

Part 2
Session 6
INCENTIVES IN SURVEYS

TIME, DOLLARS, AND DATA: SUCCEEDING WITH REMUNERATION IN HEALTH SURVEYS

Trena M. Ezzati-Rice, Andrew A. White, William D. Mosher
National Center for Health Statistics, CDC

- Maria Elena Sanchez
Agency for Health Care Policy and Research

Introduction

All surveys experience nonresponse in spite of the fact that the goal for all sample surveys is to maximize survey response, and thus minimize nonresponse bias in the survey estimates. Over the years, survey incentives (both monetary and non-monetary) have been used in conjunction with other survey methodologies to obtain complete and accurate information for the largest number of sample units.

Incentive use in surveys has spanned a wide variety of survey types, sponsors, respondents, and survey topics. As the title suggests, we present and discuss the use of remuneration (monetary incentives) in selected health surveys conducted by the National Center for Health Statistics, CDC (NCHS), the Agency for Health Care Policy and Research (AHCPR), and Project HOPE.

Experimental and field results have demonstrated that incentives can have a positive impact on survey response rates, and we will concentrate on describing selected incentive applications and experiments. Included here are new results from a field trial of remuneration in NCHS's National Survey of Family Growth, Cycle 5 Pretest.

Additional information in the area of remuneration in health surveys not reviewed in this paper can be found in the literature^{1, 2, 3, 4}. In particular, the paper by Kulka

¹ Kulka, R.A. (1992, October). A Brief Review of the Use of Monetary Incentives in Federal Statistical Surveys, presented at the COPAFS Symposium on Incentives in Surveys, Boston, MA.

² Willimack, D.K., Petrella, M., Beebe, T., and Welk, M. (1992, August). The Use of Incentives in Surveys: Annotated Bibliography, Survey Research Center, Institute for Social Research, University of Michigan.

included in this volume gives a brief background on the use of monetary incentives and references numerous articles this topic across various types of surveys. Kulka also addresses sociological models proposed to describe the operation of incentives. The paper by Willimack et al. is an unpublished review of published literature however they do make the observation : "The bulk of the published literature regards the effects of incentives in mail surveys. No doubt incentives have been and/or are being used in both telephone and face-to-face surveys, but there is a lack of documentation of tests in the published literature. Perhaps incentives have been implemented in telephone and face-to-face surveys based on the mail survey results and on 'common sense,' without specific testing within mode. Based on a brief look at conference abstracts, it appears that documentation of incentive testing and/or use in survey modes other than mail may be found in non-published literature, such as professional association conference presentations."

Overview of Health Survey Issues and Remuneration

The use of monetary incentives is not new in health surveys and over the years those of us working in the field have learned a great deal about their use. Much of the material in this section was presented at an October 1992 COPAFS/OMB Symposium on Providing Incentives to Survey Respondents held at the John F. Kennedy School of Government, Harvard University, Cambridge, MA.

Although most surveys conducted by the Federal Government are based on unpaid, voluntary participation, monetary or gift incentives for participation may be justified for certain types of household and establishment based surveys in order to increase participation rates, encourage accurate record keeping, and/or keep expenses down. In addition for health surveys, remuneration may be justified for surveys which involve a physical examination and the drawing of a blood sample in order to maximize response rates.

High response rates bring the benefits of increased validity through increased precision and reduced potential for bias in survey estimates. Incentives and

³ Kulka, R. A. (1994, May). The use of Incentives to Survey "Hard-to-Reach" Respondents: A Brief Review of Empirical Research and Current Research Practice, presented at the COPAFS Seminar on New Directions in Statistical Methodology, Bethesda, MD.

⁴ Dillman, D. A. (1991). The design and administration of mail surveys. Annual Review of Sociology, 17, p. 225-249.

remuneration can be considered appropriate whenever respondents are asked to devote time and effort (sometimes a considerable amount) to assisting the government in obtaining high quality data for research and policy related issues which will impact the entire nation.

For population based surveys, incentives can result in a higher motivation to participate, increased effort to give accurate (honest) responses, greater acceptance of government surveys, increased response rates, lower item nonresponse rates, and cost savings through better data quality and fewer call backs and/or canceled appointments. Further, cash incentives enhance the importance of the survey to respondents and provide tangible evidence of the value of their input. Incentives may stimulate otherwise reluctant respondents to participate and to more readily make themselves available to participate. Gift incentives for children could reinforce the value and importance of their participation.

The largest potential for net benefits from remuneration or incentives will be found: among surveys that experience higher refusal or item non-response rates; among surveys where persons at higher risk of being targeted respondents are less likely to participate (thus biasing the results); in situations where respondents are easy to locate, but initial or continuing cooperation is hard to gain; among those surveys requiring repeated contact in a short time period (like a month) or with other significant burden involved (like a physical examination or drawing of blood); or among those that require respondents to do something on their own, like complete a mail questionnaire or keep diaries, or participate in a followup survey of initial nonrespondents.

In particular incentives are most likely to have an effect in surveys that: require the respondent to travel; are lengthy or have a longitudinal component; are focussed at hard to reach populations (like adolescents or young black males), or that ask questions about sensitive topics (like income, drug use, risk behaviors related to HIV/AIDS). Remuneration may also gain survey participation when sample persons do not perceive an immediate benefit to themselves and/or society by participating.

Preparation and pilot testing of questionnaires (especially laboratory based testing) often requires respondent travel and takes large blocks of respondent time. The potential benefit to the survey is so large from this type of testing that remuneration is well worth it.

For institutional surveys, incentives can result in money and time savings (e.g., hospital versus government staff abstracting hospital records), removal of a barrier to participation (financial loss), greater acceptance of government surveys, cost savings

through fewer call backs and/or canceled appointments, increased response rates, and for some businesses, lower item nonresponse rates, and better data quality.

Remuneration in institutional surveys is often seen as compensation rather than as an incentive to participate in voluntary surveys. Institutions think of time as money and may consider compensation a requirement to engage in substantial continuing activities. Some may set that standard for any voluntary survey. In this sense, remuneration may make a survey possible. In addition to participation, remuneration to the institution can be cost beneficial to the government if summary data must be compiled from business records for survey purposes. The institution can often do it for less than having government field representatives comb the records. In businesses where staff energies must be diverted to complete a survey, the thoroughness and accuracy of the response may be improved if the business is being remunerated.

In institutional surveys, the continuing nature of data collection from particular respondents is a critical factor. There may be no other way to obtain the data (e.g. there is a charge for access to hospital records and state vital records already exist, why should the government be excepted?). Remuneration may actually be less expensive than providing the person hours needed to compile the data from institutional records even if the institution is willing to grant access.

A review of these issues and the use of incentive methods took place at the 1992 COPAFS/OMB Symposium⁵. This led to a recommendation, "that OMB seriously consider an agency's request to use incentives in a limited number of specific situations in which a survey violates the norm of what is considered the standard survey." The participants defined a standard survey as: a cross-sectional survey of the household population done in about an hour in a single session at the respondent's convenience; and done in the respondent's home with non-intrusive, nonsensitive questions. Symposium participants suggested that incentives be considered in the following special situations:

Surveys of hard-to-reach or special population subgroups;

Surveys involving unusual demands or respondent intrusions such as -

⁵ Council of Professional Associations on Federal Statistics (COPAFS). (1993), September). Providing incentives to survey respondents: Final report (Contract No. GS0092AEM0914). Washington, DC: Regulatory Information Service Center, General Services Administration.

Lengthy interviews,
Keeping a diary,
Taking physical or physiological tests, or
Going somewhere special to participate;

Surveys involving sensitive questions and/or topics;

Surveys involving a commitment to participate over time such as for a panel survey; and

Surveys where respondents are not a household respondent such as a physician, hospital, or nursing home.

Specific Experiences with Remuneration in Health Surveys

This section highlights some major institutional and population based surveys which successfully use or have used remuneration to increase survey response and/or data quality. For some of the summaries we have paraphrased or reproduced language in the original references for the sake of accuracy.

NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY

The National Health and Nutrition Examination Survey (NHANES) is a periodic survey designed to assess the health and nutritional status of the noninstitutional population of the United States. The NHANES consists of a household interview with adult, youth, and family medical history questionnaires, followed by a 3-4 hour standardized physical examination in specially equipped mobile examination centers (MEC's). The NHANES is based on a stratified multistage cluster probability sample design⁶. The on-going Third NHANES or NHANES III is the seventh in a series of surveys using health examination procedures that have been conducted since 1960 by NCHS.

As for most large-scale Federal surveys, the success of the NHANES surveys depends upon voluntary participation of individuals selected in the sample.

⁶ Ezzati, T., Massey, J., Waksberg, J., Chu, A., and Maurer, K. (1992). Sample Design: Third National Health and Nutrition Examination Survey. National Center for Health Statistics. Vital Health Statistics. Series 2, No. 113.

Remuneration has been used in all NHANES surveys and has been shown to be necessary for attaining adequate response rates. For the three early Health Examination Surveys (HES) conducted in the 1960's of adults, children, and youths, respectively, the examination response rates were excellent ranging from 87 percent to 96 percent⁷. However, with the beginning of NHANES I(1971-74), the examination response rate (64 percent) was much lower than those in the earlier HES surveys. After extensive efforts to improve the miserable response rate through interviewer re-training, increased publicity, and community outreach, the response rates remained at an unsatisfactory level. Therefore, it was proposed that a monetary incentive be considered to reduce examination nonresponse⁸. Since there was little information available from previous studies to show the effect of paying respondents to participate in health surveys, a field experiment was undertaken. First, of course, justification had to be submitted to OMB and approval obtained. It was reasoned that remuneration for NHANES was justified since participation in the survey required several hours of the respondent's time (thus lost time from work) and paid assistance for child care might also be required. It was hypothesized that the cost of the remuneration would be offset by a reduction in the number of contacts to a household to obtain respondent participation. Further, if the response rates increased significantly, the overall validity of the survey results would outweigh the remuneration costs.

The experimental design for the study was superimposed upon the within primary sampling unit (PSU) design for NHANES I⁹. The study was undertaken during 1972 in the San Antonio, Texas, PSU. The segments within the PSU were randomly paired by segment size and median family income. All of the sample persons in one segment of each pair were told about the \$10 remuneration. The sample persons in the other segment of the pair were not told of the \$10 remuneration. It should be noted however, that all persons who were examined received \$10. The difference was that persons in the "not told" segments did not know about the remuneration until after they had been examined, while those in the "told" segments knew of the \$10 remuneration before being examined.

Telling sample persons that they would be given \$10 after completing the examination phase of the survey had a positive effect on the response rate in San

⁷ Bryant, E., Kovar, MG., and Miller, H. (1975). A Study of the Effect of Remuneration Upon Response in the Health and Nutrition Examination Survey. National Center for Health Statistics. Vital Health Statistics. Series 2, No. 67.

⁸ Ibid.

⁹ Ibid.

Antonio. Among the 303 persons in the experimental segments (told) who were contacted, 82 percent were examined (see Table 1). On the other hand, among the 292 persons in the control segment (not told), only 70 percent were examined. Thus, the NHANES I experiment showed that the offer of \$10 to sample person increased the response rate by 12 percentage points. Also, there was some evidence that sample persons were more cooperative and that less effort was required to obtain response when remuneration was offered as evidenced by the number of persons making an examination appointment at the first interview contact. Also, a larger proportion of the "told" group kept their appointments than the "not told" group. Only 2.1 contacts per examined person were required for the "told" group as compared with 2.5 such contacts per examined person for the "not told" group.

Table 1. Examination Response Rates from NHANES I Remuneration Experiment (sample sizes shown in parentheses)

	Not Told of Payment	Told of Payment
Experiment	70%	82%
	(292)	(303)
Actual Survey*	68%	78%
	(7335)	(6035)

*Represents results from the first 35 stands of NHANES I.

The findings of the NHANES I remuneration study were considered conclusive enough to include remuneration in the remainder of NHANES I. The overall response rate at the 45 survey locations where remuneration was offered in NHANES I (including San Antonio) was 78 percent as compared to 68 percent for the 19 survey locations where remuneration was not offered.

The \$10 remuneration used in NHANES I was continued in NHANES II. However, about midway through NHANES II, the response rate dropped to about 70 percent. Therefore, another field experiment was used to assess the effect of increased remuneration on survey response. The study for NHANES II (1976-80), looked at the impact of increasing the \$10 remuneration used in NHANES I to \$20. Three survey locations were selected for the study. Each segment was paired with another segment similar to it with regard to poverty/non-poverty status and distance from the

examination center. Sample persons in one of each pair of segments were "told" that they would receive \$20 for participating in the examination, while the sample persons in the other paired segment were told they would receive \$10. However, all persons who were examined were given \$20 no matter which monetary amount they were originally told.

There was a significant positive effect on response rates with the \$20 incentive¹⁰. Out of 720 persons offered \$20 for participation, 79 percent were examined (see Table 2), while among the 716 persons offered \$10, only 74 percent were examined. An important finding from NHANES II which did not show up for NHANES I was the increase in the response rate by number of persons in a household for the \$20 group versus the \$10 group. There was no effect in households with only one sample person. However, in households with two and three or more sample persons, payment of \$20 rather than \$10 increased response rates by 8 and 16 percentage points, respectively¹¹.

Table 2. Examination Response Rates from NHANES II Remuneration Experiment (sample sizes shown in parentheses)

	Told of \$10 Payment	Told of \$20 Payment
Experiment	74%	79%
	(716)	(720)
Actual	72%*	76%**
	(NA)	(NA)

*Represents the final 44 stands of NHANES II.

**Represents the first 16 stands of NHANES II.

¹⁰ Findlay, J. and Schaible, W.L. (1980) A Study of the Effect of Increased Remuneration on Response in a Health and Nutrition Examination Survey. Proceedings of the Section on Survey Research Methods of the American Statistical Association, pp. 590-594.

¹¹ Ibid.

The results indicated that there was a potential cost saving associated with the fact that the higher paid group was generally more cooperative. For example, it was found that a larger proportion of the examined respondents in the higher paid groups went to the examination center as a result of the first contact (68 percent versus 61 percent). Also, the higher paid group required, on average, fewer contacts overall (1.77 vs. 2.09)¹².

Overall, for NHANES II, for the 44 survey locations at which \$10 was offered to sample persons, the response rate was 72 percent. But for the 16 stands at which \$20 was given, the response rate was 76 percent.

In the Puerto Rican phase of Hispanic HANES (1984), the response rates were unsatisfactory, so an increase in remuneration from \$20 to \$50 was made. A cross-tabulation of response rates (see Table 3) by method of payment before and after the increased remuneration was instituted indicated a higher response rate in the \$50 group (83 percent) than for the \$20 group (76 percent)¹³. The results, however, should be viewed with caution since the increased payment was not randomized within the various survey locations.

Table 3. Hispanic HANES Examination Response Rates for New York City Metro Area (Puerto Rican Phase) Stands by Payment Amount*, 1984

Payment Amount	Examination Rate
\$20	76% (3101)
\$50	83% (576)
DK**	0% (116)

*Unpublished data from J. Findlay, NCHS.

**Cases could not be classified into either payment category due to lack of information.

¹² Ibid.

¹³ Unpublished data from Jean Findlay, National Center for Health Statistics.

In the on-going NHANES III, the basic remuneration payment is \$30 for all persons who come to the examination. However, for adults 20 years and older there is an additional incentive of \$20 if they come for their examination at the "right time". Time of day of the examination and fasting status need to be controlled for the analyses of many of the biochemical tests including the Oral Glucose Tolerance Test (OGTT - a test for diabetes). Thus a random half-sample of adults is designated to have blood drawn in the morning, while the remaining 50-percent is selected for afternoon or evening appointments.

The NHANES III which includes an oversample of both Mexican-Americans and Blacks and includes no upper age limit (in contrast to previous NHANES) is experiencing higher examination response rates than any other previous NHANES survey. There is a six percentage difference in the overall examination response rate between NHANES II and NHANES III-Phase 2 (79% in the on-going NHANES III-Phase II versus 73% in NHANES II - see Table 4). This increase is due not only to the increased remuneration amount but also to the fact that NHANES III samples multiple persons per household and includes on average two persons per household. Clearly, there is a strong monetary incentive for a household as a whole if several members are selected into the sample.

Table 4. Examination response Rates for NHANES II, Hispanic HANES, and NHANES III

	MEC Examined	MEC + Home Examined
NHANES I	74	NA
NHANES II	73	NA
HHANES	73	NA
NHANES III - PHASE 1	77	78
NHANES III - PHASE 2*	79	80

*On-going.

NHANES III examination of nonresponse rates by household size show that the nonresponse rates decrease significantly with increasing household size. Another important finding related to participation in NHANES III is the increase in the

response rate for the OGTT. This test requires the sample person to fast 10-16 hours and requires two blood draws. The OGTT response rate in NHANES II among examined persons was only 66 percent thus requiring the results to be interpreted with extreme caution. However, for NHANES III, the OGTT response rate, again among examined persons, has increased by 11 percentage points (77 percent response rate for phase 1) and by 23 percentage points so far for phase 2 (89 percent response rate)¹⁴. The increased remuneration in NHANES III is felt to be responsible in large part for this important increase in the OGTT response rate for NHANES III.

NHANES Summary

Previous research associated with NHANES I and NHANES II has involved several field experiments to assess the impact of monetary incentives on achieving an adequate response rate so reliable estimates can be produced and generalized to the total population. The NHANES I remuneration experiment generally showed that paying sample persons \$10 rather than nothing at all increased response rates about 12 percent.

The NHANES II study looked at the effect on the response rate of increasing the remuneration from the \$10 used in NHANES I to \$20. The average increase in the response rate was about 5 percent.

The results from both the NHANES I and NHANES II had two other findings pertinent to remuneration. First, the results indicated that there were potential cost savings associated with the fact that the higher paid groups were generally more cooperative. For example, it was found that a larger proportion of the examined respondents in the higher paid groups went to the MEC as a result of the first contact. They also required, on the average, fewer contacts overall. They also had fewer broken appointments.

Further, both of these experimental studies showed a marked increase in cooperation in households with more than one sample person, and the differences between the remuneration groups went up with household size. A positive relationship between household size and response rate was also observed in the Hispanic HANES. This trend is continuing in NHANES III as well. These results, presumably, are due to the fact that the monetary award increases substantially for the household as a whole when multiple persons are selected per household. Finally, the response rates in

¹⁴ Unpublished data from Meena Khare, National Center for Health Statistics.

NHANES III are higher than those for previous NHANES surveys and the response rates for the two minority subgroups are higher than for whites and all others. In addition, the response rate for the OGTT in NHANES III is higher than for NHANES II.

The NHANES experiences demonstrate the power of remuneration to stimulate respondent participation in the face of inconvenience and financial disincentives; and in addition to motivate respondents to provide highly personal and sensitive information including physical and biological characteristics. The use of remuneration in NHANES III shows how phased payments can improve scheduling logistics as well. The remuneration serves as a motivator, and also to offset financial disincentives, and gives evidence of the importance of participation and survey goals. NHANES data have been significantly improved through incentive use.

NHANES I EPIDEMIOLOGIC FOLLOWUP STUDY

As mentioned in the previous section, examinees were paid \$10 to participate in the baseline NHANES I study. They were again paid \$10 to participate in the first NHANES Epidemiologic Followup Study (1982-84 NHEFS interview). Although no remuneration was paid in the 1986 followup of the elderly (conducted using computer assisted telephone interviews: CATI), subjects with reported hospital or nursing home admissions in the 1987 Followup were paid \$5 as an incentive to sign and return the Medical Authorization Form (MAF). Payment was needed in the 1987 Followup because of the sharp decrease in the rates at which MAF's were returned in the 1986 Followup. Remuneration of the 1987 respondents increased the MAF return rate in the elderly group by 11 percentage points from 75.0% in the 1986 Followup to 85.7% in the 1987 Followup.

For the 1992 wave, a \$5 remuneration was made subject to reported hospital or nursing home admissions since the last contact. Respondents are interviewed by telephone. However, the form which authorizes the hospital or nursing home to release patient information is mailed to the respondent and must be signed and returned. The remuneration is paid to the respondent or to the individual who signs the MAF, if this person is different from the respondent, as an incentive to return the signed authorization.

In addition, due to the difficulty of gaining the cooperation of some hospitals and nursing homes, if a nursing home or hospital requests reimbursement for work performed in abstracting or photocopying selected information from the admission and discharge records or abstracts, a small amount of money is provided to defray their

expenses. During 1987, a total cost of \$2,170 or .0012 percent of the total contract cost was paid. This remuneration mostly involved nursing homes.

The NHEFS uses remuneration to gain commitment to continuing participation in a longitudinal study.

NATIONAL HEALTH INTERVIEW SURVEY YOUTH RISK BEHAVIOR SUPPLEMENT

In 1991 the Survey Research Center of the University of Michigan and the Bureau of the Census helped the National Center for Health Statistics assess the impact of financial rewards on respondent participation and motivation in a Youth Risk Behavior Supplement (YRBS) to the 1992 National Health Interview Survey (NHIS). This study¹⁵ employed cognitive interview techniques and a traditional split sample field experiment.

The YRBS contained a number of sensitive questions requiring answers that could be perceived as being socially undesirable or threatening to self-image. Such topics included questions on the use of alcohol, drugs, sexual activity, over- or underweight status, and exercise program involvement.

Interviews were conducted both with groups of parents and youth before field trials and a pretest were conducted. The field trials were then held where youth were interviewed in their homes, paid \$20, and debriefed after the interview about their views on paying respondents for participating in surveys. The pretest of the full survey procedures included a split sample in which the interviewers mentioned the \$20 payment half of the respondents and did not mention the payment to the other half.

In the split sample experiment, complete interviews were obtained in 90% of the households where the \$20 was mentioned, but in only 79% of the remaining households. The cognitive interviews conducted with the groups and during the field trial debriefings helped the researchers develop insight concerning reasons for this difference in response and for a potential increase in data quality when payment is

¹⁵ Kalton, G., Cannell, C., Camburn, D., Oksenburg, L. and Holland, L. (1991). The Effect of Financial Incentives on Respondent Participation. Final report of the Association of Schools of Public Health Cooperative Agreement: "Applied Research on the Conduct of Adolescent Health Behaviors and Characteristics", University of Michigan, pp. 47-56.

mentioned in advance.

The cognitive interviews indicated that unwillingness to report events or behaviors is only partially caused by concerns over privacy or confidentiality. In fact, youth respondents could see no reason to be diligent in answering survey questions. They voiced skepticism about any benefits that may accrue from survey participation. Therefore techniques typically used with adult respondents, such as appealing to their obligations as citizens or emphasizing the societal benefits may be unproductive. The researchers concluded that a \$20 remuneration offer, linked with a signed commitment to participate and give accurate answers, would be an effective way to motivate respondents to participate and report accurately. The debriefing interviews confirmed this conclusion.

A further interesting note from this study is that the success of respondent payments in obtaining YRBS interviews may be partly due to the reaction of interviewers to the payments. During interviewer debriefings, their comments suggested that respondent payments also have a forceful, positive influence on the attitudes and expectations of interviewers. Interviewers with such a positive outlook may feel they are likely to obtain an interview, rather than expecting a refusal, prior to contacting a potential respondent and subconsciously may convey to potential respondents a more positive view of the YRBS study. The researchers postulate that the total impact of respondent payments on participation rates is the sum of the positive direct effect on respondents and the indirect effect that payments have on the attitudes of interviewers.

YRBS incentive use focused on the motivational aspects of incentives for a non-traditional target population in a study that collected sensitive data.

NATIONAL HOUSEHOLD HIV SEROPREVALENCE SURVEY FEASIBILITY STUDY

A feasibility study for a National Household HIV Seroprevalence Survey (NHSS), based on a probability sample of households was conducted in Dallas County, Texas, in the fall of 1989. One of the major concerns of a household survey attempting to estimate the prevalence of HIV infection is that a high proportion of persons who are at the greatest risk of HIV infection may refuse to participate. This possibility of differential rates of response between those at higher risk and lower risk of HIV infection means that estimates derived from the survey have the potential to be biased. Among procedures to maximize the response rate in the NHSS, a \$50 incentive payment was provided to all sample persons who provided a blood sample to be tested

for HIV antibodies and completed a self-administered risk behavior questionnaire. The \$50 incentive seemed to have a positive impact on survey participation with a higher than expected response rate for a highly sensitive survey involving the collection of HIV risk behavior data and the collection of a blood specimen in the home. The response rate for the combined questionnaire and blood sample was 84 percent, and 90 percent for the questionnaire only¹⁶. Respondents in the NHSS were asked to check all reasons for their participation in the survey. Among all respondents, 47 percent stated "helping with AIDS research", while another 39 percent stated the "\$50 payment" (see Table 5).

Table 5. Reasons for participation in the Dallas County Household HIV Survey

Reasons	Percent
Helping with AIDS research	46.8
U.S. Public Health Service Sponsorship	2.9
\$50 payment	38.5
Videotape presentation	1.2
Assurance of privacy	3.3
Other factors, unspecified	7.2

In addition to estimating the prevalence of HIV infection in Dallas County, another major objective of the survey was to evaluate various methods for assessing and reducing nonresponse bias. A standard survey method for assessing bias due to nonresponse is to conduct a followup survey with a sample of initial survey nonrespondents with different incentives for participation. In the NHSS, a special followup study of a sample of nonrespondents was conducted in which half of the sample persons were offered an incentive of \$100 to complete the self-administered risk behavior questionnaire only, and the other half was offered \$175 to complete both the questionnaire and provide a blood sample. The followup survey of nonrespondents increased the questionnaire only response rate by 10 percentage points (80% vs. 90%)

¹⁶ National Household Seroprevalence Survey Feasibility Study Final Report. Research Triangle Park, NC: Research Triangle Institute, April 30, 1990. Research Triangle Report RTI/4190-01F.

and the blood and questionnaire by four percentage points (80% vs. 84% - see Table 6). Of particular importance was the increased reporting of risk behavior in the follow-up study. The prevalence rates among male respondents for three major HIV risk behaviors (intravenous drug use, receptive anal intercourse, and multiple sex partners) were 3 to 5 times higher in the followup survey than in the regular survey (see Table 7).

In summary, results from the NHSS feasibility study followup survey indicated that a high proportion of persons who initially refused to participate, when recontacted and offered an increased incentive, completed the risk behavior questionnaire. A lower proportion of persons who initially refused to participate provided both a blood sample and completed the risk questionnaire, when recontacted and offered an increased incentive. Persons at higher risk for HIV infection participated at higher levels in the followup survey than in the regular survey. The followup survey effectively increased the total number of persons who participated in the Dallas HIV survey. The increase in risk reporting among the sample of regular survey nonrespondents that were followed-up allowed for a significant reduction in nonresponse bias in the HIV estimate produced for Dallas County.

Table 6. Sample persons response rates in the Dallas County Household HIV Survey, 1989

Survey component	Regular survey	Regular + followup survey	Overall*
Screening	97	98	98
Blood & Questionnaire	80	84	82
Questionnaire only	80	90	88

*Product of screening and sample person rate.

Remuneration in the NHSS feasibility study demonstrated the power of incentives even in the most sensitive topic and invasive data collection situations, however establishing the exact mechanics of the reasons for success would require further study.

Table 7. Prevalence of selected HIV risk behaviors in the regular and followup survey, Dallas County male population, 18-54 years, 1989

Risk Behavior Since 1978	Regular Survey	Followup Survey	Total Estimate
Intravenous drug use	3	12	4
Receptive anal intercourse	3	11	5
5+ male partners	2	10	3
1+ male partner	5	17	8

THE NATIONAL SURVEY OF FAMILY GROWTH PRETEST

The National Survey of Family Growth (NSFG) is done periodically by NCHS to collect national data on the factors that affect the U.S. birth rate and women's reproductive health--factors that include sexual activity, marriage and divorce, contraception, sterilization, infertility, miscarriage, and abortion. Previous cycles of the NSFG have interviewed about 8,000 women 15-44 years of age in the noninstitutional population of the United States with response rates ranging from 75 percent to 80 percent.

Interviewing for the next NSFG, called Cycle 5, will be conducted in January-July of 1995. Three of the principal challenges for Cycle 5 of the NSFG will be (1) increasing response rates to make it possible to conduct a telephone reinterview in 1997 with as many of the original respondents as possible, (2) improving reporting of HIV-related sexual behavior, and (3) improving the reporting of abortion.

Response rates.--Most recently, in 1988, the NSFG used a list sample of households interviewed in the NHIS. Using a list sample saves nearly a million dollars on sample design and selection costs, but it makes it necessary to find women who move between the NHIS interview and the NSFG interview. Some are never found, so response rates are reduced somewhat. Response rates have been between 75 and 80 percent in recent cycles, despite the intrinsic appeal of the subject matter, the use of only female interviewers, thorough interviewer training, advance letters

introducing the survey, and expensive, intensive nonresponse follow-ups¹⁷.

Although the NCHS obtains data from the National Health Interview Survey (NHIS) on HIV-related knowledge and attitudes, the NSFG remains the NCHS's principal vehicle for collecting data on HIV-related behavior, including such sensitive topics as age at first intercourse, numbers and characteristics of sexual partners, and condom use.

Abortion reporting is critical in the NSFG Cycle 5 because 25% of all pregnancies, and half of all unintended pregnancies, end in abortion. Fertility surveys in the U.S. and other nations have obtained incomplete reporting of abortion. In the last 3 cycles of the NSFG, in 1976, 1982 and 1988, and in most other U.S. surveys, women reported less than half of the abortions they have actually had¹⁸. This incomplete reporting of abortion has several potential adverse effects: it makes impossible analyses of the determinants and consequences of abortion itself; it forces us to use ad hoc methods to produce estimates of pregnancy rates for the U.S.; it produces biased estimates of the failure rates of contraceptive methods¹⁹; and it forces us to study unintended births instead of unintended pregnancies.

The NSFG Pretest for Cycle 5, conducted in October-December 1993, was based on about 800 eligible women, of whom 500 completed interviews. The Pretest was an experiment, which was designed to test several alternative contexts for asking questions. The pretest had 3 main groups:

- 1) the first group was a standard Computer-Assisted Personal Interview in the respondent's home with no incentive.
- 2) in the second group, the interview was moved to a neutral site--a site outside the home--where spouses, children, or parents could not hear the respondent's answers. To reimburse women for the time and inconvenience of going to the neutral site, respondents were paid \$40 in cash at the end of the interview.
- 3) when we considered these first two groups, we were concerned that we might

¹⁷ Judkins, David P, W. Mosher, and S. Botman. (1991). National Survey of Family Growth: Design, Estimation, and Inference. Vital and Health Statistics, Series 2, No. 109.

¹⁸ Jones, Elise and J. Forrest. (1992). Under-reporting of Abortion in Surveys of U.S. Women: 1976-1988. Demography 29 (1): 113-126.

¹⁹ Jones, Elise and J. Forrest. (1992). Contraceptive Failure Rates Based on the 1988 NSFG. Family Planning Perspectives 24 (1): 12-19.

obtain higher response rates and data quality in the neutral site/\$40 group than in the in-home no-incentive group, but we would not know whether the differences were due to the non-home site or to the \$40 payment. We were also concerned that the costs of implementing a \$40 payment and setting up non-home sites on a national scale might be prohibitive. Therefore, we added a third group--a \$20 incentive for an interview in the home.

For half the respondents in group 1 and group 3, (the in-home interviews), we also tested a short questionnaire at the end of the interview, using Audio CASI (Computer Assisted Self-Interviewing with headphones, and the respondent entering her answers into the computer).

NSFG Results

Pretest response rates (as a percent of those located) were higher for incentive cases than for non-incentive cases: 81 percent for those who received a \$20 incentive vs 73% for those who received no incentive. The percent who broke an appointment with the interviewer was one-third lower for those who received \$20 than for those who received no incentive (24 vs 37%).

The number of hours that the average interviewer worked to get a completed case was about 2 hours less for incentive than for non-incentive cases (8.8 vs 10.9; note also that when the incentive increases to \$40, hours per case drops more than 2 full hours - see Table 8). Since the time of interviewers costs more than \$10 an hour for their wages plus benefits, if the interviewer can save 2 hours of effort per case by paying a \$20 incentive, then the incentive pays for itself. That is precisely what happened in the NSFG Pretest: the incentives paid for themselves in the \$20 group because respondents broke fewer appointments for interviews and made themselves available after fewer telephone calls and personal visits. Costs in the \$40 non-home group were higher because of high costs to set up the neutral sites--obtaining permissions, renting office space, etc.

Reporting of the number of sexual partners was higher among respondents who received incentives, still lower than reports of comparable studies of men^{20, 21}. In

²⁰ Billy, J.O.G., Tanfer, K., Grady, W.R., and Klepinger, D. (1993). The Sexual Behavior of Men in the United States. Family Planning Perspectives 25 (2): p. 52-60.

previous cycles of the NSFG, less than half of abortions were reported²². Thus, abortion reporting should be at least double what it was in the last Cycle. In the pretest, a \$20 incentive and the use of Audio CASI doubled abortion reporting, compared to a no-incentive, no Audio CASI group. This more complete reporting of abortions is probably due to two factors: (1) women using the Audio CASI (headphone) questionnaire reported a higher percentage of their abortions, and (2) incentives produced better coverage of groups of women who have higher abortion rates--including black women and poor women.

Table 8. Response Rates, Broken Appointment Rates, and Interviewer Hours per Case from the NSFG Pretest for Cycle 5

Experimental Group	Response Rate (%)	Percent with Broken Appointment	Hours per Case
In-Home, No Incentive	73	37	10.9
In Home, \$20	81	24	8.8
Non-Home, \$40	80	31	6.4

Recommendations from the Pretest

The results from the NSFG pretest suggest that a \$20 incentive plus Audio CASI (self-administered questionnaires over headphones) should be used in the NSFG main study. The incentive will increase response rates, particularly among minorities and low-income women, and reduce the cost of interviewer labor because respondents will cooperate more readily.

Can these results be generalized? The results on response rates, interviewer hours and costs, in the NSFG Pretest are quite similar to those in the National Adult Literacy

²¹ Smith, T.W. (1991). Adult Sexual Behavior in 1989: Number of partners, frequency of intercourse and Risk of AIDS. Family Planning Perspectives 23 (3): p. 102-107.

²² Op cit. Jones and Forrest, Demography, Feb. 1992.

Survey (NALS) Field Test, a survey with a much larger sample (n=2,000) than the NSFG Pretest. Like the NSFG, the NALS required considerable effort from respondents²³.

The NALS and NSFG Pretest results provide evidence that incentives may be most cost-effective when the interview is:

(1) either long or a great deal of effort, or both; (2) sensitive either because it deals with private behaviors or may otherwise cause embarrassment (the NALS might cause such embarrassment among the adult illiterate, the NSFG because it includes questions on abortion and sexual behavior); and (3) part of a panel survey in which the response rate is critical to maintain the size of the panel over time.

All three of those conditions were common to both the NALS and the NSFG. The NSFG experience clearly demonstrates the success of incentives with hard-to-interview populations and sensitive topics. In addition, it clearly demonstrates the cost-effectiveness of the methodology in improving survey quality.

NATIONAL MEDICAL EXPENDITURE SURVEYS

The National Medical Expenditure Surveys (NMES) are designed to produce estimates of medical use, medical expenditures, sources of payment for medical care, and health insurance coverage. The surveys are sponsored by the Agency for Health Care Policy and Research (AHCPR). The household survey (HS) component of the NMES series yields estimates for persons in the civilian non-institutionalized population, while the survey of the institutionalized population (IPC) yields estimates for persons in nursing homes. Incentives have not been a design feature of the IPC surveys; therefore, the remarks that follow will focus on the use of incentives in the NMES HS surveys.

Respondent Incentives

The NMES series of household surveys includes the 1977 National Medical Care Expenditure Survey (NMCES), the 1987 National Medical Expenditure Survey (NMES2), and the 1996 National Medical Expenditure Survey (NMES3) presently in

²³ Berlin, M, L. Mohadjer, J. Waksberg, et al. (1993). An Experiment in Monetary Incentives, in American Statistical Association (editor), 1992 Proceedings of the Section on Survey Research Methods, pages 393-398.

the planning phase. These NMES studies share the following design features:

- the use of an initial screening interview to identify the households to be sampled for the study;
- the oversampling of poor people, blacks, Hispanics, and the elderly;
- a longitudinal or panel design in which sampled families are interviewed several times over a period of 14-16 months, in rounds of data collection that cover a year-long observation period;
- face-to-face interviews lasting on average between 2 and 2.5 hours in each round to be completed with a family respondent that provides information about him/herself and all other family members;
- the request that respondents prepare for the interviews by keeping a study calendar and saving records such as bills and insurance statements in order to improve the accuracy of their reporting of medical use and expenditures;
- special requests for information in addition to the interview itself, such as completion of self-administered forms;
- requests for signed permission forms from specific sample persons authorizing the collection of data from medical providers, employers and other health insurance providers to supplement and validate the data obtained from households.

In the context of the core NMES design summarized above, respondent incentives have been used primarily for the following reasons:

- 1) to motivate respondents to participate initially and in future interviews in order to minimize initial nonresponse and panel attrition.
- 2) to compensate respondents fairly for the burden associated with long interviews and the completion of additional survey tasks; and
- 3) to motivate respondents to keep records and provide fair compensation for the effort required to maintain the study calendar and save financial records over a long period of time.

The respondent incentives most frequently used in NMES surveys consist of cash

payments made by the interviewer at the end of each interview, typically in the form of a check. Checks provide a safe and convenient mode for interviewers to handle cash payments. In NMES respondents were paid \$5 at the end of each interview, beginning with Round 1. The amount was increased to \$10 per interview in NMES2, and to a proposed \$15 per completed interview in NMES3.

NMES respondents are required to sign a receipt acknowledging that money was received. In NMES2, the receipt was used as a vehicle to consolidate commitment to the role of respondent in the survey. The receipt the respondent was asked to sign included a statement indicating the willingness to accept responsibility for record-keeping in preparation for the next interview.

The use of incentives has no doubt contributed to the high response rates achieved in NMES studies, in spite of the burden that long interviews represent, and notwithstanding the oversampling of groups that in many surveys yield lower than average response rates, such as poor people and minorities. The overall response rate for the NMES2 household sample was 80.1 percent after four rounds of data collection.

Incentives are an important tool used by interviewers to convince reluctant respondents to participate in NMES surveys. Recent methodological research examined the characteristics of persons who had initially refused to be interviewed in any one of the NMES2 rounds of data collection²⁴. The analysis revealed that reluctant respondents differed significantly from their cooperative counterparts with regard to the proportion of overall medical expenditures that different sources of payment covered for each group. The reduction of significant differentials with respect to health expenditures and insurance coverage, two core analytic concerns of NMES surveys, provide evidence of the beneficial use of incentives to guard off against potential nonresponse bias in national estimates.

In 1985, to aid in planning for NMES2, a feasibility study²⁵ was conducted to investigate a broad range of methodological issues. Among the issues examined was

²⁴ Cohen, S.B. and B.L. Carlson (1992). "An Analysis of the Characteristics of Reluctant Respondents in the National Medical Expenditure Survey". Proceedings of the Section on Social Statistics, American Statistical Association, in press.

²⁵ Mathiowetz, N.A. and Ward, E.P. (1987). Linking the National Medical Expenditure Survey with the national Health Interview Survey: An Analysis of Field Trials, Vital and Health Statistics, Series 2, No. 102, U.S. Government Printing Office, Washington, D.C.

the effect of several incentive protocols²⁶. The feasibility study consisted of two rounds of data collection: 1) a personal interview 45-60 minutes long where the respondent was paid five dollars at the end of the interview, and 2) a second interview in person or by telephone where again the respondent was paid five dollars at the completion of the interview. Approximately two weeks prior to the second interview, self-administered questionnaires were mailed to all respondents. These questionnaires were designed to take approximately 30 minutes to complete and included some moderately threatening questions on health behavior and mental health status.

Reporting groups were divided into three treatment groups: 1) "Prepayment" - persons were sent a five dollar check with the questionnaire; 2) "Promised" - persons were told that they would be paid five dollars when the completed forms were returned; and 3) "No mention" - persons were not given any information on payment, but were paid five dollars upon questionnaire completion.

The results indicated that the prepaid incentive leads to a significant improvement in response rates. Seventy-three percent of those in the prepaid group completed the survey compared to 66% among those who were not told of the incentive. The prepaid incentive also worked better than the promised incentive, which resulted in a response rate of 60%. Item nonresponse rates were calculated for each completed self-administered questionnaire and were used as a general measure of data quality. The finding was that prepayment leads to lower item nonresponse. Ninety percent of those who were prepaid answered all of the questions in the 18 page questionnaire, compared to only 74% in the promised group and 87% in the no mention group.

The study concluded that prepaid incentives can result in higher response rates and more complete data with less need for follow-ups. These were achieved in this study at a very moderate increase in cost. However, the net added costs may be far less than the value of the incentive payments, since a substantial part of the incentive costs is offset by savings in the follow-up activities. This finding was in line with results from other mail surveys, and that mode of payment was adopted in NMES2 when self-administered forms were mailed to sample households.

Based on results from the most recent NMES feasibility study carried out in 1992, the schedule of cash incentives in NMES3 will be modified relative to earlier surveys. Instead of introducing record-keeping tasks and paying respondents for the first time at

²⁶ Berk, M.L., Mathiowetz, N.A., Ward, E.P. and White, A.A. (1987). The Effect of Prepaid and Promised Incentives: Results of a Controlled Experiment. Journal of Official Statistics, Vol. 3, No. 4, pp. 449-457.

the end of Round 1, the 1992 Feasibility Study presented the family respondent with the study calendar at the end of the Screener round and prepaid these respondents in anticipation of the time and effort that would be devoted to keeping records in order to prepare for the Round 1 interview. Payments for later rounds of the Feasibility study were also described as compensation for future effort. In the last round a token gift (a commemorative tile with the U.S. Public Health Service seal) was given to respondents instead of cash payment. Under the new plan, the gift is the only net increase in incentive costs compared to previous plans, and it appears that prepayment has advantages. The rates of NMES calendar use in Round 1 of the Feasibility Study are significantly higher than the rates achieved in NMES2 in the round immediately following the round when payment and instructions to keep records were first delivered²⁷.

The 1992 NMES Feasibility Study also tested successfully the use of incentives to motivate respondents to complete a complex data collection task that was time dependent. This involved the procurement of health insurance printed materials that included a description of the benefits associated with the health plan offered by the employer to each policyholder in the household²⁸. In NMES2 the collection of comparable information was attempted from employers in the course of the health insurance provider survey, but the time lag between the end of the household survey and the start of that provider survey frequently made it impossible to locate the necessary information about the health plan in effect at the time the household was interviewed.

In the Feasibility Study, respondents were offered \$15 per household to contact employers, either by mail or in person, and secure the necessary information. Interviewers gave respondents a request form that could be presented to the employer to facilitate the task. Payment was made when the health insurance materials were delivered in the next round, and the amount was not increased in the event that the family had more than one eligible policyholder. Health insurance booklet requests in the Feasibility Study were followed up in later rounds and, by the end of the study, a policy booklet had been retrieved for 75 percent of eligible plans, at a lower cost and

²⁷ Sanchez, M. E. (1993). "Enhancing Compliance with Record-keeping in a Household Survey". 1993 Proceedings of the Section on Survey Research Methods of the American Statistical Association. Vol. 2, p 1015-1020.

²⁸ Emmons, C.A., Curno, M., and Smith, K. (1993). Final Report on the Outcomes of the Procedure for Obtaining Health Insurance Policy Documents from Respondents in the NMES3 Household Survey Feasibility Study. Submitted by NORC and Westat, Inc. under contract requirements.

on a more timely basis than in NMES2.

The use of remuneration to improve survey scheduling, achieve self-response and record keeping, and establish a commitment for a longitudinal data collection process are shown in the NMES experience. The NMES experience also demonstrates the potential for improved data quality with remuneration.

Interviewer Incentives

The morale and motivation of interviewers and supervisors are important factors in response rate outcomes for sample surveys. While the impact of respondent incentives has been frequently discussed in the survey research literature, there is little evidence of systematic inquiry into the use of incentives for interviewers as a means of achieving high response rates.

Typically, studies resort to interviewer incentives in a haphazard and improvised fashion when production levels have tumbled and the study response rate is deemed unacceptably low. This plan of action may not be as desirable or as economical as the notion of setting up a planned and carefully crafted incentive plan for interviewers from the very start of the project. The experience in the 1992 NMES Feasibility Study with such a plan suggests the desirability of exploring further the manner in which interviewer incentives may be manipulated to achieve gains in field response rates and efficiency within acceptable budget limits.

The Feasibility Study included a plan for interviewer incentives in order to achieve high response rates within the schedule for data collection. The project staff and the contractor collaborated to come up with a plan that was acceptable to all. The field staff wanted a plan that would promote and reward team work as opposed to individual performance; thus, the team was defined as the cluster of interviewers working in each PSU.

Realistic response rate levels for two points in time during the round (a specified midpoint in the field period, and the end of the round) were defined for each PSU and communicated to interviewers at the beginning of each round. The incentives were cash payments (about \$25 for each of the time points in a round) paid equally to all PSU interviewers provided the PSU had achieved the targeted response rate by the specified date. A very modest additional amount of money was paid for increases in the response rate beyond the specified minimal rate.

The bonus plan encouraged interviewers to talk to their fellow PSU interviewers to

coordinate the steady flow of work. A weekly memo informed interviewers of the progress in other areas and the ranking of their own PSU. With very few exceptions, the staff in each of the PSUs achieved the response rate goals consistently for a very modest investment.

The benefits of the incentives included: teamwork among the staff at the local PSU level; steady production which avoided the last minute dislocations and expenses typically associated with a late push to increase response rates; and availability of qualified staff towards the end of the study to handle difficult assignments in a planned fashion.

More systematic research on the use of interviewer bonuses and the performance of different bonus plans is desirable, as the strategy appears a cost-effective way of obtaining high response rates.

Remuneration can be a valuable tool for direct interviewer management as well as helping interviewers motivate respondents to participate.

Physicians are often surveyed to obtain a wide variety of medical information. The high frequency at which physicians are surveyed coupled with the problem of "gate keepers" probably contribute to the low response rates typically achieved in physician surveys. However, prepaid incentives seem especially effective for this population group^{29, 30, 31, 32}. The following three experiences serve as examples.

²⁹ Berry, S. H. and Kanouse, D.E. (1987). Physicians response to a mailed survey: An experiment in timing of payment. Public Opinion Quarterly, 51, p. 102-114.

³⁰ Lockhart, D.C. (1991). Mailed surveys to physicians: The effect of incentives and length on the return rate. Journal of Pharmaceutical Marketing & Management, 6(1), p. 107-121.

³¹ Mizes, J.S., Fleece, E.L., and Roos, C. (1984). Incentives for increasing return rates: Magnitude levels, response bias, and format. Public Opinion Quarterly, 48, p. 794-800.

³² Berk, M.L., Edwards, S.E. and Gay, N.L. (1993). The use of a prepaid incentive to convert nonresponders on a survey of physicians. Evaluation & The Health Professions, Vol. 16, No. 2, pp.239-245.

NATIONAL SURVEY OF DIAGNOSTIC ALLERGY TESTING

In 1988 a remuneration experiment was conducted³³ on a subset of physicians selected to participate in the National Survey of Diagnostic Allergy Testing (sponsored by the Health Industry Manufacturers Association).

A sample of physicians was randomly divided into three experimental groups. The first group received a \$10 incentive with the first mailing. Nonresponders to the initial mailing were sent a new questionnaire as well as a letter urging them to respond and mentioning the \$10 incentive they had received earlier. The second group of physicians did not receive a monetary incentive with the initial mailing. On the second mailing of the questionnaire, however, they received another letter explaining the importance of the study as well as a \$10 prepaid monetary incentive. No mention of payment was made to the third group on either the first or second mailing.

The results indicate the use of a prepaid incentive has a dramatic impact on the response to the initial mailing. Fifty-five percent of those physicians receiving a prepaid incentive responded to the initial mailing, compared to less than 20% who were not told about payment on the initial mailing. Overall, a 63% response rate was obtained for Group 1 physicians (prepaid incentive with the initial mailing), compared with only a 50% response rate for Group 2 (prepaid incentive on the first prompt), and a 40% rate for Group 3 (no incentive).

This study concluded that incentives should be used in cases in which its use is considered necessary to obtain adequate response rates. Also, while delaying the decision to use an incentive until the second wave of mailing enables the researcher to decide whether an adequate response rate is likely to be obtained without payment, the incentive is not nearly as effective when used in a follow-up mailing. Because obtaining high response rates on physician surveys is difficult, few researchers will be able to conclude at study onset that a high response rate can be obtained. The use of a prepaid monetary incentive enclosed with the initial questionnaire mailing, therefore, appears to be a cost-effective method for improving response rates on physician surveys.

NATIONAL HOSPITAL DISCHARGE SURVEY

Most of this paper has dealt with remuneration for individual respondents, but

³³ Op cit. Berk et al., Evaluation & the Health Professions, 1993

institutions and organizations are sometimes respondents for surveys, and the National Hospital Discharge Survey (NHDS) experience suggests that remuneration is essential in some institutional surveys.

The National Hospital Discharge Survey (NHDS), conducted by NCHS, was first fielded in 1964 following the completion of a feasibility study³⁴. The NHDS is a continuing study designed to provide comprehensive general-purpose statistics on morbidity in patients discharged from the Nation's general and short-stay hospitals. The principal source of information for the survey is the medical record in the hospital. The data are obtained from probability samples of medical records abstracted in a sample of general hospitals. Hospitals are compensated for participation in the survey.

Some hospitals in the sample have automated records, and contract with an abstract service. Other hospitals abstract records manually as needed. There are two manual data collection procedures: a primary manual procedure in which hospital staff complete the abstracts, and an alternate procedure in which a Census Bureau representative completes the abstracts. Hospitals using the primary manual procedure receive an average of \$2.40 per abstract submitted; those using the alternate manual procedure receive about \$1.00 per abstract submitted. Data tapes of uniform abstracts covering all discharges for automated hospitals are purchased directly from abstract service organizations. The cost of these data ranges from \$.003 to \$.055 per discharged patient. These discharges are sampled for the survey. All hospitals participating in the NHDS are reimbursed \$1.00 per record biannually for approximately 40 records that are reabstracted for quality control procedures.

A large part of the success of the NHDS depends on the willingness of the hospitals to perform substantial continuing activities. Once inducted into the survey, hospitals participate for an extended period of years. A substantial amount of work is involved, including sampling the discharge lists, pulling and refiling medical records, and abstracting approximately 20 records monthly. A feasibility study conducted in 1963-1964 found that most hospitals expected compensation for their effort.

Cost analysis supports the practice of remuneration, particularly for primary manual procedure hospitals. For example in Fiscal Year 1992, it cost the NHDS an additional average of \$12.15 per record to have the Census Bureau sample and abstract the data

³⁴ Brown, A.M., Altman, I. and Thompson, D.J. (1966), Participation of Hospitals in the Pilot Study of the Hospital Discharge Survey, Vital and Health Statistics, Series 2, No. 19, U.S. Government Printing Office. Washington, D.C.

in the alternate procedure hospitals. If the Bureau of the Census performed this work in all the sample hospitals, the cost of the NHDS would be substantially increased.

The feasibility study found that "While some hospitals indicated they might be willing to collaborate in the survey without reimbursement, it was clear that most would expect some compensation for their contributions, especially where it was felt that overtime work or the employment of additional personnel might be required. The ... form shown to the [hospital] administrator ... was quite detailed and implied a fairly exhaustive review of the medical record. It contained questions on the characteristics of the patient and his *sic* hospitalization, including final diagnoses, operations, complications, laboratory tests, therapies, and the like. There were some differences about whether payment should be made to the hospital or to the personnel doing the work, but most administrators favored payment to the hospital."

The feasibility study recommended that "... a uniform policy be adopted for the compensation of hospitals and that fair payment, based on further examination of the true cost to the hospitals be made." A subsequent pilot study was used to confirm the acceptability of the survey procedures, including remuneration, and helped to calibrate the payment amounts.

The NHDS experience illustrates the testing and use of remuneration and their use to offset a financial disincentive, improve the acceptance of a government survey, provide evidence of the value of participation, and keep expenses down in a survey requiring the continuing participation of institutions.

Summary and Conclusions

The evidence summarized in this paper shows that remuneration for respondents can be an effective technique for raising response rates and data quality when otherwise good survey practices are not sufficient.

More and more in the last decade, policy makers and program administrators are demanding data that cannot be supplied with a standard survey -- the one-hour cross-sectional interview containing nonsensitive questions is no longer the norm. The health surveys reviewed here each have one or more features that do not fit the public's perception of what a standard survey is: some require long interviews (NSFG Pretest and NMES); others require the maintenance of records such as diaries (NMES) or event histories (NSFG); some are panel surveys with repeated interviews (NSFG, NHANES Follow-up, NMES); some have sensitive questions (NSFG Pretest, NHSS, and YRBS); others use non-home sites (NHANES and the NSFG Pretest); or ask for

medical tests (NHANES); or require information or testing that could be embarrassing (NALS, NHANES, and NHSS).

As interviews get longer and questionnaire content gets more difficult or intrusive, and hard to interview sub-populations are surveyed, the need to motivate respondent participation grows. We need to provide respondents with concrete evidence of our appreciation and the importance of their participation and willingness to provide accurate and complete information. The careful use of remuneration allows us to offer people more than the promise that policy makers including Congress will use the data to improve their lives. If an advance letter or first personal contact explains that they will be compensated, all of the practical evidence reviewed here suggests that completing a quality interview at a reasonable cost is more likely.

Although most of this paper has dealt with remuneration for individual respondents in households, it is important to note that institutions, medical professionals, and other organizations are frequently respondents for health surveys. The NHDS experience suggests that remuneration is important in some institutional surveys also.

The history of remuneration in health surveys as evidenced by the experience reviewed here is a successful one. Remuneration has stood the test of time, proving successful in controlled experiments, field trials, and long-term implementation. For relatively little cost, important improvements in response and data quality have been gained using remuneration methods.

THE USE OF INCENTIVES TO SURVEY "HARD-TO-REACH" RESPONDENTS: A
BRIEF REVIEW OF EMPIRICAL RESEARCH AND CURRENT RESEARCH PRACTICE

Richard A. Kulka
National Opinion Research Center
University of Chicago

1. Introduction

Incentive payments to survey respondents have been used extensively for many years as a means of improving survey response rates (cf. Shuttleworth, 1931), and there is considerable research evidence supporting the value of compensation for increasing cooperation and improving the speed and quality of response in a broad range of data collection efforts (cf. Kulka, 1992; Willimack, Petrella, Beebe, and Welk, 1992). In particular, a large number of empirical studies concerned with increasing response to mail questionnaires consistently attests to the effectiveness of monetary incentives in increasing mail survey response rates (e.g., Armstrong, 1975; Church, 1993; Duncan, 1979; Fox, Crask, and Kim, 1988; Harvey, 1987; Heberlein and Baumgartner, 1978; Hopkins and Gullickson, 1992; Kanuck and Berenson, 1975; Linsky, 1975; Yammarino, Skinner, and Childers, 1991; Yu and Cooper, 1983).

Nevertheless, until fairly recently monetary incentives and other forms of respondent remuneration have not been used extensively in general survey practice, especially in studies under government sponsorship and large scale academic surveys--as opposed to their widespread and common use in commercial or market research. In recent years, however, it has become increasingly difficult to achieve response rates high enough to provide statistically valid results, and remuneration has become more common. In the United States, Federal statistical surveys cannot employ incentives without explicit authorization from the Office of Management and Budget (OMB), and OMB has generally prohibited the use of payments to respondents, except under circumstances where "substantial need" can be demonstrated. And, interpreting this rule on a case-by-case basis--as requests for the use of incentives from Federal agencies have become increasingly common--has made it more difficult for OMB to assure consistency in the application of these guidelines.

To assist OMB in developing appropriate principles and decision rules governing the use of respondent incentives by Federal agencies, in October, 1992, the Council of Professional Associations on Federal Statistics (COPAFS) convened a symposium of representatives of government, business, academic, and research organizations to consider the current state of experience, research, knowledge, and opinion regarding the use of such incentives (COPAFS, 1993). Although it was not expected that the symposium would provide definitive answers to the multitude of

questions surrounding the use of incentives, "OMB expected to obtain information that would help in preparing guidelines to foster greater consistency in reviewing future requests by Federal agencies to use incentives when conducting surveys" (p. 1).

Toward that goal, symposium participants discussed, in part, "the kinds of survey situations in which incentives have a high probability of being effective or necessary" (p. 8), articulating "a set of circumstances in which they thought OMB should seriously consider an agency's request to use incentives" (p. 9). Among these were a number of situations or circumstances that might be broadly conceived as under the general rubric of "hard-to-reach respondents." Specifically, the list included the use of incentives:

- To encourage *hard core refusals* to respond, especially in small subpopulations of interest where response rates are often quite low--low enough to raise serious questions about the quality of survey data for these subpopulations.
- When there is a significant likelihood that a "gatekeeper" will prevent the respondent from ever receiving the questionnaire or otherwise make it difficult to make contact with certain segments of the population to conduct an interview.
- When there is a special target population for whom our conventional means of motivation or encouragement will have little if any chance of working--i.e., where the positive forces to cooperate are quite low (e.g., prostitutes, the homeless, the disenfranchised).
- If the target population is a small group that is often surveyed, such that a particular respondent is likely to be sampled frequently for one survey or another (e.g., physicians, CEO's, university deans).
- If the population is a *control group* in a program evaluation or experiment in which it is imperative to achieve and maintain an adequate response rate in the control group sample if the integrity of the study is to be maintained.

More generally, one of five potential OMB incentive policies suggested for consideration in nonstandard survey situations was that incentives "be considered if the respondent incurred out-of-pocket cost; or if the survey was too intrusive; or the survey was aimed at a *hard-to-reach population* [emphasis added]" (COPAFS, 1993, p. 12). Overall, most of those present felt that:

- hard-to-reach really meant *hard-to-interview* [emphasis added]. This category could include those who are hard to encourage to cooperate, and therefore initially refuse. In such cases, incentives might be effective. [However], participants felt that incentives would not be effective for those who are hard to find.

Participants also included in the hard-to-interview category those who are difficult to reach by mail, those who must be kept in a sample (such as members of a control group), and those disenfranchised from society. (p. 12).

Based on this broad conception, the focus of this paper is explicitly on the use of incentives to survey hard-to-reach respondents, in contrast, for example, to the use of respondent incentives as a reimbursement for out-of-pocket expenses, as a payment to respondents for their time and effort in participating in a survey, or to compensate respondents for carrying out survey tasks that entail unusual demands, i.e., those which are especially burdensome or intrusive or may put the respondent at risk. In principle, to the extent that "hard-to-reach" is viewed as synonymous with "hard-to-survey," this focus is quite broad, in that encouraging those who might otherwise be reluctant to cooperate with a given survey to indeed do so is, in essence, the basic intent of the vast majority of surveys that choose to provide remuneration to survey respondents as part of their design.

However, incentives provided to stimulate survey response are rarely given only to initial or hard-core refusals, and several important questions regarding the use of incentives to encourage response may be addressed by focusing on this particular use of remuneration (cf. COPAFS, 1993). These include:

- Are there indeed specific target populations who are routinely offered remuneration to participate in surveys by most or all survey organizations because they are regarded as especially difficult to survey?
- While respondent incentives may increase cooperation among initial refusals, are they really effective with hard-core refusals and the truly difficult or impossible to interview populations?
- Are incentives effective only for certain target populations or subpopulations or more effective for certain population subgroups than for others (i.e., are the effects of incentives different for different population subgroups)?
- Are incentives really effective in getting past "gatekeepers," either for certain professionals (e.g.,

physicians) or other difficult-to-survey subgroups of the general population?

- Are incentives indeed of little use in locating and interviewing hard-to-find cases?
- Should consideration be given to paying some, but not all respondents to a given survey?
- Should all respondents be paid the same incentive, or should consideration be given to different levels or types of remuneration for different respondents?

To seek possible answers to these and some other related questions, we conducted a focused review of the current research literature on the use of incentives, with particular attention to their use with hard-to-reach populations, broadly defined.

2. Current Research Practice

However, because we anticipated that empirical evidence bearing on many of these questions would likely be quite sparse, we also sought input from individuals at most of the government, business, academic, and research organizations represented at the COPAFS symposium in October 1992, along with a few others, to ascertain the current state of survey research practice with regard to the use of incentives to survey hard-to-reach populations. In addition to providing citations or references to any papers or publications related to this topic, each organization was asked to provide information on any recent surveys they had conducted with such populations, indicating when they had or had not used incentives, and a sense of their general organizational policies or conventions regarding the use of incentives under such circumstances. Prior to describing some of the evidence available from the research literature that bears on the questions raised above, it will be useful to summarize current practice in this area as described by these organizations.

Not surprising, the vast majority of these organizations routinely conduct surveys with hard-to-reach respondents under our broad definition. As noted by one organizational respondent, virtually every survey encounters and must deal with hard core refusals, but more specific categories of respondents designated by responding organizations as hard-to-reach are:

- (1) the economically disadvantaged (e.g., lower income or lower socioeconomic status (SES) populations, welfare recipients or applicants, the homeless);

- (2) the educationally disadvantaged (e.g., the less educated, high school dropouts, those with low literacy levels, the mentally retarded);
- (3) minority populations (e.g., African Americans, Hispanics, disadvantaged minorities, impoverished urban minorities);
- (4) adolescents, youth, and young adults (e.g., youth in general, minority youth, young black males, teen mothers, the young and mobile);
- (5) drug users and those with special health problems (e.g., current or former drug users, drug abusers, cocaine users, diabetics, those with asthma);
- (6) frequently surveyed professional or elite populations (e.g., physicians, nurses, CEO's, teachers, college and university faculty, both very small and very large farm operations); and
- (7) transients and persons who wish not to be found for legal or other reasons (e.g., highly mobile and transient populations, runaway youth, absent parents owing child support, those defaulting on student loans).

Not only do these categories overlap a great deal, but also, in almost every case, incentives have been used in surveys of these populations to increase response rates by at least one organization, and quite often by many. For example, the use of (generally substantial) incentives in surveys of physicians is a standard practice in virtually all of these organizations. It is also the case, however, that recent surveys have been conducted with most of these subpopulations in which no incentives were used, including a few surveys of physicians.

Although most of the organizations queried feel that respondent incentives are generally effective in increasing response rates among these hard-to-reach target populations, very few controlled or randomized experiments have been conducted to demonstrate empirically the efficacy of incentives in improving response rates under such circumstances. Not surprisingly, the use of incentives in surveys conducted by or for Federal statistical agencies--which require OMB approval for providing respondent incentives--is somewhat more likely to be based on such empirical evidence than their use in surveys conducted by commercial, academic or private research firms under other auspices.

Even in the absence of such controlled experiments, several of these survey research professionals and firms believe that respondent incentives are an important overall tool in their arsenal for dealing with hard-to-interview survey populations, and

their experiences with such incentives bear on at least three of the basic issues raised earlier regarding the use of incentives with such populations. First, although many participants at the COPAFS (1993) symposium felt that incentives might not be effective in locating those who are hard-to-find, several of these organizations report experiences that suggest that paying respondents makes contacting and locating easier, less expensive and more effective, since contact individuals are more willing to convey messages and provide new address and telephone numbers for sample members when interviewers are able to mention that they have a monetary incentive for the respondent. Similarly, others cite experiences suggesting the efficacy of respondent incentives in "opening the door" or getting past "gatekeepers,"--i.e., in helping gain access to the respondent--because nurses, receptionists, relatives, friends and other "gatekeepers" are apparently more reluctant to restrict or deny access to a potential respondent when a monetary incentive is involved.

Third, although several of those responding to our inquiry expressed some ambivalence regarding this practice, a number of surveys have been conducted which provide incentives either to some but not all respondents, or different levels of remuneration to respondents in the same survey. In some cases, these different incentive levels (including no incentive) reflect different levels of burden, effort or risk for different respondents, but it is also not uncommon to offer incentives only to sample members for whom one is having difficulty getting them to respond--i.e., to do the survey without routine remuneration, but then use monetary incentives to try to convert refusals. Alternatively, in a survey providing incentives at a given level, interviewers might be permitted to offer increasingly larger amounts to convert increasingly hard-core refusals, to persuade extremely hard-to-convince cases to indeed cooperate.

Although these conditional incentive approaches can be quite cost effective, paying uncooperative sample members when cooperative respondents are not paid, or paying especially reluctant or difficult sample members more than those who cooperate more readily, violates our sense of fairness or equity. However, the selective or strategic use of remuneration to convert hard-core refusals and achieve higher or very high response rates is--though relatively rare and practiced with some reluctance--very much a part of current research practice with regard to the use of incentives to survey hard-to-reach respondents.

3. A Brief Review of the Literature

Having briefly summarized the current state of practice in this area, let us now explore what, if anything, the current research literature can tell us regarding the use of respondent incentives with difficult-to-survey populations or sample members.

3.1 The Use of Incentives in Surveys--A Summary

As a background for that analysis, it is important to consider momentarily what we know (or think that we know) about the use of respondent incentives in general (cf. Kulka, 1992, Willimack et al. 1992). As noted earlier, few today would question the general assertion that a monetary incentive enclosed with a mail questionnaire will serve to increase response rates. Hundreds of studies have been conducted, and review after review--both qualitative and quantitative--concludes that the importance of financial incentives is "second only" (perhaps) to the use of follow-up mailings or prompts in improving response rates (cf. Dillman, 1991). Moreover, the literature rather overwhelmingly supports the predominant effectiveness of prepaid as opposed to promised incentives (e.g., Armstrong, 1975; Berk, Mathiowetz, Ward and White, 1987; Berry and Kanouse, 1987; Blumberg, Fuller, and Hare, 1974; Church, 1993; Furse and Stewart, 1982; Hopkins and Gullickson, 1992; James and Bolstein, 1992; Kanuck and Berenson, 1975; Linsky, 1975; Peck and Dresch, 1981; Skinner, Ferrell, and Pride, 1984; Wotruba, 1966; Yu and Cooper, 1983). That is, incentives appear to be most effective in inducing survey response when they are paid in advance--at the time that the respondent's cooperation is initially solicited--rather than offered conditional on and paid subsequent to respondent cooperation, even when the promised or conditional incentive is greater than the amount prepaid (cf. Linsky, 1975; James and Bolstein, 1992).

The use of monetary incentives to increase response rates for telephone and personal interview surveys has received far less research attention, although the results of several studies are consistent with those derived from mail surveys. Overall, however, the conditions under which a monetary incentive will be effective or ineffective under these survey modes appear to be less general. Based on this more limited research literature, the greatest potential effectiveness of monetary incentives appears to be in surveys that place unusual demands upon the respondent, require continued cooperation over an extended period of time, or when the positive forces on respondents to cooperate are fairly low (cf. Cannell and Fowler, 1977).

In addition to a potential beneficial impact on response rates, the research literature suggests that incentives may have a beneficial impact on data quality as well. At least two theories suggest the opposite--a detrimental effect of remuneration on data quality:

- (1) a concept based on "social desirability" theory (Cannell and Henson, 1974; Weiss, 1975) that suggests monetary inducements will increase the tendency of participants to try to please the interviewer by providing what the respondent believes is the desired or "correct" answer; and

- (2) a "self-perception" model that argues that the introduction of financial incentives acts as an external motivator, thereby reducing the degree of internal motivation (i.e., the subject's interest or desire to participate), decreasing the degree or quality of compliance (i.e., quality of response), while increasing the rate of compliance or cooperation (cf. Hansen, 1980).

While relatively little empirical evidence has been found in support of either of these two models (see, however, Hansen, 1980; James and Bolstein, 1990), the preponderance of evidence reported to date (e.g., Berk et al., 1987; Cowan, 1977; Ferber and Sudman, 1974; Godwin, 1979; Goetz, Tyler, and Cook, 1984; Houston and Ford, 1976; James and Bolstein, 1990; Kerachsky and Mallar, 1981; McDaniel and Rao, 1980; Sudman and Ferber, 1974) is more consistent with a theory based on "social exchange," which posits that the offer of monetary incentives induces a greater commitment to the survey task among respondents, which in turn results in better data quality from survey respondents.

With regard to incentive size, the research literature is significantly less helpful, since the majority of studies have investigated the effects of incentives of \$1 or less (e.g., Armstrong, 1975; Fox et al., 1988; James and Bolstein, 1992; Jobber and Saunders, 1988; Kanuck and Berenson, 1975; Linsky, 1975; Yammarino et al., 1991; Yu and Cooper, 1983), and few studies have successfully demonstrated the effectiveness of very large monetary incentives (e.g., Berry and Kanouse, 1987; Gunn and Rhodes, 1981; James and Bolstein, 1992). At both extremes, there is some evidence that increasing the size of monetary incentive will result in increases in survey response and/or quality (e.g., Armstrong, 1975; Findlay and Schaible, 1980; Fox et al., 1988; Furse and Stewart, 1982; Godwin, 1979; Gunn and Rhodes, 1981; Hubbard and Little, 1988a, 1988b; James and Bolstein, 1990, 1992), but there is also rather consistent evidence that this benefit may rather quickly reach "diminishing returns," whereby larger incentives no longer result in appreciable increases in survey response (e.g., Armstrong, 1974; Fox et al., 1988; Godwin, 1979; Hubbard and Little, 1988b; James and Bolstein, 1992; Mizes, Fleece, and Roos, 1984).

Overall, why payments are effective in improving response rates is not currently very well understood. Some believe that "payment works in increasing response rates . . . through motivating and supporting the interviewer [emphasis added] in his [or her] approach to the respondent" (Weinberg, 1973, p. 480), while others view "incentives . . . as impressing upon the survey respondent [emphasis added] the importance of the task" (Goetz et al. 1984:149; Berry and Kanouse, 1987). Still others have suggested that the predominant motivating power of an incentive is not its monetary value, but rather its symbolic, or "token" value (e.g., Linsky, 1975).

In turn, two basic classes of theories have been referenced to explain why incentives may increase survey participation (cf. Willimack et al., 1992). The first is based on the principle of "reciprocation" (Groves, Cialdini, and Couper, 1992):

Every human society abides by a *norm of reciprocity* [emphasis added] that directs individuals to provide to others the general form of behavior that they have received from others (Gouldner, 1960). . . . [Based on] the reciprocity heuristic, . . . one should be more willing to comply with a request to the extent that the compliance constitutes the repayment of a perceived gift, favor, or concession. (p.480)

Thus, by providing an incentive as an unsolicited gift (e.g., a prepaid incentive), one invokes the norm of reciprocity among respondents, who can "reciprocate" by participating in the survey.

Closely related to the concept of a "norm of reciprocity" (and in the same basic class) are theories of "cognitive dissonance" and "social exchange." Under the former, it is postulated that the inclusion of a prepaid token incentive with a request for survey participation creates psychological dissonance, which is most easily resolved by consenting to in fact participate (cf. Furse and Stewart, 1982, 1984; Hackler and Bourgette, 1973).

As articulated by Dillman (1978), "social exchange" theory, which emphasizes the perceived costs and rewards of responding to a survey, suggests that, in order to maximize survey response, one must "minimize the costs for responding, maximize the rewards for doing so, and establish trust that those rewards will be delivered" (Dillman, 1978, p. 12). Rather than serving as a reward for survey participation, the use of an incentive serves as "a symbol of trust," a major factor necessary for social exchange to successfully occur. Consistent with this notion is research evidence showing that

increasing the size of an incentive does not always increase response, and in fact may tend to decrease it, and that including it with the appeal [a prepaid incentive] is more effective than promising to send it on return of the questionnaire. The closer the monetary incentive comes to the value of the service performed, the more the transaction tends to move into the realm of economic exchange [in which money serves as a precise measure of the worth or value of one's actions] and the easier it becomes for people to refuse it. (Dillman, 1978, p. 16)

In general, smaller, prepaid incentives appear to invoke social exchange or the norm of reciprocity, while larger promised or conditional incentives are more likely to invoke economic exchange, which represents the second basic class of theories on how incentives serve to increase survey participation, i.e., by

literally paying respondents for the time and effort required to provide the information requested. Under this model, remuneration represents reimbursement for survey cooperation (cf. Cohen, Walden, and Ward, 1992), i.e., compensation to respondents for their time and effort in participating in the survey, rather than an "incentive," "gift," or "gesture of goodwill" too small to be considered equitable payment or exchange for respondents' time.

In this regard, Cannell and Henson (1974) emphasize that, since respondents rarely share the goals of a survey, they do not consider participation as a means of advancing their own, personal goals, and are thereby generally unmotivated to perform the survey task. An incentive, in the form of a payment, may serve to provide a personal goal which motivates participation in a survey. To be effective, the amount of money offered must be large enough to be worth working for--i.e., the value of the incentive should be commensurate with the task and/or time sacrificed by the respondent--but not extravagant, because, if individuals perceive that they are overcompensated, the effects on participation may actually be negative (Cannell and Henson, 1974; Groves, 1989).

Overall, the preponderance of research evidence appears to favor the concepts of social exchange or reciprocity as a basis for the effectiveness of incentives, although with the advent of increasingly more complex surveys and the use of larger incentives, a significant body of evidence consistent with the tenets of economic exchange is also accumulating, much of it in relation to surveys of apparently hard-to-reach or hard-to-interview populations, to which we now turn our full attention.

3.2 Hard-to-Reach Target Populations

Although at least seven different categories of respondents were identified by survey practitioners as "hard-to-reach," research on the effects of respondent incentives has focused directly on only a few of these--frequently surveyed professional and elite populations, adolescents and young adults, and the disadvantaged--and the relative coverage of even these three broad populations in that regard is quite uneven.

By "focusing directly" I mean that the subjects for experimentation with incentives are drawn largely or entirely from one of these target populations. A number of these studies include one or more of these subgroups as a component of the population surveyed, and such subgroups may or may not have responded differently to incentives than other groups represented in the sample, but the issue of the differential effectiveness of incentives among those hard-to-survey will be dealt with in a separate section.

Studies on the use of incentives with professional and elite populations--especially physicians--are legion. Physicians are

widely believed to be an especially difficult population from which to collect survey data (Sudman, 1985), since they "are frequently approached for surveys, the demands on their time are great, and their office staffs are vigilant in protecting them" (Berry and Kanouse, 1987, pp. 102-103). Physicians' reluctance to participate in surveys is a growing problem for researchers (cf. Berk, 1985), with the American Medical Association (AMA) warning that "physicians are becoming weary and wary of surveys" (Martin, 1984), while the response rates to even their own surveys have declined precipitously (Goodman and Jensen, 1981).

Both prepaid (Berk, Edwards, and Gay, 1993; Berry and Kanouse, 1987; Lockhart, 1991; Mizes, Fleece, and Roos, 1984) and promised (Gunn and Rhodes, 1981; Weber, Wycoff, and Adamson, 1982; Tambor, Chase, Faden, Geller, Hofman, and Holtzman, 1993) monetary incentives have been shown to significantly improve response rates in surveys of physicians, whether conducted by mail, telephone or in person. For example, Gunn and Rhodes (1981) conducted an experiment to determine the effectiveness of paying monetary incentives to physicians for their participation in a 20-30 minute telephone interview on attitudes toward influenza immunization. Physicians were systematically assigned to one of three subsamples designated to receive no incentive, \$25, or \$50. Study findings revealed a 58 percent response rate for the group offered no incentive, 69 percent for those promised \$25, and 77 percent for those offered \$50. In a personal interview survey conducted by Weber and his colleagues (1982) the same range of incentive conditions resulted in response rates of 38, 67, and 73 percent, respectively.

Berry and Kanouse (1987) compared the relative effectiveness of a prepaid and a promised incentive of \$20 in a mail survey of physicians, obtaining a 78 percent response rate for the prepaid incentive group and a 66 percent rate for those paid only after they completed the survey. Mizes, Fleece, and Roos (1984) demonstrated the effectiveness of even a relatively small prepaid incentive in a brief mail survey of physicians, obtaining a response rate of 74 percent with either a \$1 or \$5 prepayment in comparison with 53 percent when no payment was provided. In a survey of 600 physicians from three specialty groups, Lockhart (1991) achieved a 57 percent response rate using a \$20 prepaid incentive, compared with only 13 percent in a no incentive control group. More recently, Berk, Edwards and Gay (1993) confirmed the effectiveness of a prepaid incentive of \$10 in a mail survey of physicians, achieving a response rate of 63 percent for those receiving the incentive with an initial mailing in comparison with 40 percent for the no incentive group.

The relative effectiveness of both large and small monetary incentives in improving response rates, speed, and/or quality among

other professional and elite populations has also been demonstrated empirically, including:

- (a) international elites, ranging from university professors to cabinet ministers (Godwin, 1979);
- (b) nurses (e.g., Kephart and Bressler, 1958);
- (c) librarians (Hopkins, Hopkins, and Schon, 1988);
- (d) various professionals subscribing to a magazine dealing with alcohol and drug use problems (Goodstadt, Chung, Kronitz, and Cook, 1977);
- (e) owners of small construction subcontracting companies (James and Bolstein, 1992);
- (f) community elites (Paolillo and Lorenzi, 1984);
- (g) business executives (Erdos and Morgan, 1983; Robin and Walters, 1976); and
- (h) farmers (Willimack, 1993).

However, some of these studies were poorly designed and/or obtained very low response rates even with incentives, and there are other studies where incentives used with professionals were either ineffective (e.g., Cook, Schoeps, and Kim, 1985; Abraham and Johnson, 1993) or resulted in poorer response rates or quality than when no incentives were provided (e.g., Hansen, 1980). For example, in the 1992 field test for a national survey of college and university faculty (Abraham and Johnson, 1993), three incentive conditions were used, including one monetary (a prepaid \$2 bill); only the monetary incentive approached statistical significance in improving the response rate over no incentive (87 versus 79 percent), and this was not regarded as strong evidence of the efficacy of incentives with that particular professional population. In fact, the main survey, fielded without incentives, achieved an overall response rate of 87 percent (Abraham, 1994).

A second category of respondents which many nominate as hard-to-reach are adolescents, youth, and young adults. While there is less empirical evidence available with regard to these target populations, the results that are available are consistent with the assumption that incentives can be quite effective in stimulating survey cooperation among them. For example, a recent investigation conducted for NCHS by the Survey Research Center at the University of Michigan (Cannell and Camburn, 1991) studied the effects of respondent payments of \$20 on the willingness of youth 12-20 to participate in the Youth Risk Behavior Surveillance System (YRBS) and on their motivation to answer YRBS questions as accurately and truthfully as possible. The results of this research indicated that paying respondents increased participation rates (from 79 to 90 percent), reduced parental consent refusal rates (thereby assisting in getting by an important "gatekeeper"), aided

interviewers in converting refusals, increased respondents' perceptions of the importance of the survey, and may have improved the accuracy and honesty of responses.

The authors suggest that these youth represent the classic hard-to-interview respondent, in that the positive forces for them to respond and respond accurately to the YRBS are in fact quite low. "Young people had a low interest in this survey, and saw no compelling reason for responding honestly or being diligent in the task of answering the survey questions" (Cannell and Camburn, 1991, p. 1). Moreover,

respondents could see no compelling reason to be diligent in answering survey questions. Participants in group interviews voiced skepticism about any benefits that may accrue from participating in surveys. Therefore, techniques typically used with adult respondents to encourage participation and more accurate reporting, such as appealing to their obligations as citizens [civic duty] or emphasizing the societal benefits likely to come from the survey, may be unproductive with youthful respondents. Therefore, some different ways of motivating respondent participation and accurate reporting are needed. (Cannell and Camburn, 1991, p. 11)

This is consistent, of course, with Cannell and Henson's (1974) earlier suggestion that this is precisely the role that a respondent incentive might play, i.e., providing a personal goal which motivates participation in a survey.

Similarly, students and former students have traditionally been difficult to survey. For example, "postsecondary students who have little motivation to participate in . . . research and have traditionally been difficult to survey include those who fail to complete the application process, those who are admitted but do not register for classes, dropouts, and alumni" (Zusman and Duby, 1987, p. 73). Thus, in a mail survey of undergraduate transfer students who subsequently withdrew, Zusman and Duby (1987) found that a prepaid incentive of \$1 increased cooperation by nearly 20 percentage points. In an earlier study of the use of incentives with a postsecondary student survey, Huck and Gleason (1974) found that the response rate could be increased from 65 to 92 percent with the provision of an incentive. Peck and Dresch (1981) found that a prepaid \$3 incentive with a 1½ hour mail survey of men and women three years after they completed high school yielded a response rate of 76 percent, compared with a 68 percent response rate for those promised a similar payment, and 54 percent among those who were offered no payment at all. Similarly, in a one-year follow-up mail survey of vocational-technical school graduates, Pucel, Nelson, and Wheeler (1971) found that the use of multiple nonmonetary incentives increased response rates by over 20 percentage points relative to a no incentive control group.

In contrast, in the 1992 Postcensal pretest for the NSF's 1993 National Survey of College Graduates, a mail survey with telephone and in person follow-ups, Mooney, Giesbrecht, and Shettle (1993) found that a \$5 incentive with the initial mailing significantly increased response rates after two mailings by 11 percentage points, but, after mail and telephone follow-up, this difference was reduced to only 2 percent. In addition to increasing speed of response, however, those provided an incentive were significantly more willing to provide telephone numbers and names of contact persons, thereby reducing the effort required for future locating in this longitudinal study. In yet another incentive experiment conducted in connection with the National Assessment of Educational Progress (NAEP) household survey -- in which young adults 26-35 were asked to complete a series of tests -- Chromy and Horvitz (1978) demonstrated the effectiveness of a variable incentive procedure (no incentive for one package, \$10 for two, \$15 for three, four for \$20) in increasing the overall response rate from 70.5 percent (with no incentive) to 83.3 percent, a rate which was maintained in subsequent years by adopting this procedure.

Perhaps the respondent categories most commonly mentioned as hard-to-reach or hard-to-interview are the educationally or economically disadvantaged and minorities. However, there is very little experimental evidence available attesting to the efficacy of respondent incentives based specifically on these target populations. One of the classic examples in fact overlaps our previous category, dealing as it does with disadvantaged youth. In that study (Kerachsky and Mallar, 1981), a \$5 payment per interview was offered to a randomly selected portion of a national probability sample panel of 5,800 economically disadvantaged youth for three waves of interviewing, conducted in conjunction with an evaluation of the economic impact of the Job Corps program. "Youth in the age range of Corpsmembers (16 to 21) and with their economically disadvantaged backgrounds are generally very mobile and hard to locate [and interview]" (Kerachsky and Mallar, 1981, p. 263).

The researchers were able to verify the effectiveness of these respondent payments for improving both the quantity (search efficacy and interview completions) and quality (e.g., item nonresponse) of responses. After one or two interviews, the study's ability to locate potential respondents and obtain data from those who were located deteriorated in the absence of monetary incentives, but not when sample members were offered the \$5 payment per interview. In addition, payment influenced the willingness of sample members to return postcards from advance letters, thereby reducing the cost of locating respondents for follow-up interviews. Significant reductions in item nonresponse associated with these payments were most pronounced at baseline and declined over time. More generally, the overall effectiveness of respondent incentives in increasing response rates and quality in studies that

overrepresent the poor and minority populations has also been demonstrated (e.g., Berk et al., 1987).

While studies of the effectiveness of incentives that focus directly on low income, minority, and disadvantaged populations are quite rare, those that address the more specific question of whether respondent incentives are more effective in improving cooperation among such target populations than the more affluent and advantaged are considerably more common. This question is part of a more general one--are incentives more effective for certain population subgroups than for others?--to which we now turn our attention.

3.3 Differential Effects of Incentives by Target Population

Restating our original question on this issue:

Are incentives effective only for certain target populations or subpopulations or are they more effective for certain population subgroups than for others (i.e., are the effects of incentives different for different population subgroups)?

At the most general level, a few of the quantitative reviews of the research literature on techniques designed to enhance mail survey response rates have addressed this issue. Yammarino and his colleagues (1991) found, for example, that associations between incentives and response rates "were statistically significant but derived from more than one population; that is, the relationships are situation specific and there is a need to examine potential moderators" (p. 627), other than year of publication and (a crude measure of the) type of sample (consumer versus institutional groups), the two moderators included in the study, which were not statistically significant. In contrast, neither Church (1993) nor Hopkins and Gullickson (1992) found significant differences in the effectiveness of incentives by target population (general, students, technical, business, and medical) or population type (professional, general, or semiprofessional), respectively.

However, a more detailed examination of respondent populations reveals a significant amount of variation in responsiveness to respondent incentives. For example, in an experiment recently reported by Willimack (1993) in connection with the 1992 Farm Costs and Returns Survey (FCRS), a prepaid nonmonetary incentive increased the overall response rate from 58 to 63 percent, but the incentive proved to be most effective in the smallest and largest classes of farm operations, increases of 17 and 12 percentage points, respectively.

Similarly, several studies of physicians have found significant differences by specialty in the effectiveness of incentives in increasing survey cooperation. Noting that different populations are more receptive to certain incentives than others, Erdos and Morgan (1983) reported that doctors in nine specialties

of medicine responded well to a 25¢ incentive, whereas psychiatrists required a \$1 incentive. In the study cited earlier by Gunn and Rhodes (1981), response rates varied considerably across specialties: pediatricians and industrial physicians had high rates regardless of payment category (\$0-\$25-\$50), but general and family physicians were very sensitive to payment--37 percent responded with no payment and 64 percent with \$50. In the Lockhart (1991) study, the impact of incentives on general and family practitioners and physicians specializing in internal medicine was substantially greater than that observed among diabetes specialists, presumably because the topic of the survey--blood glucose monitoring--was especially salient to the latter. In contrast, in Berry and Kanouse's (1987) investigation of prepaid versus promised incentives, prepayment was effective with all nine specialties (statistically significant in four), except for oncologists, whose overall response rate was exceptionally high.

While it is commonly assumed that monetary incentives are more effective in increasing response rates among less educated, lower income respondents than among more educated, middle- or upper-income persons or households, research findings are in fact somewhat mixed regarding the effect of incentives on response rates for groups differing in socioeconomic status. For example, based on the early mail survey literature, Kanuk and Berenson (1975) cite several studies that failed to show that low income people were more likely to respond to monetary incentives than were people with higher incomes. In the 1971 National Health and Nutrition Examination Survey (NHANES I; Miller, Kennedy, and Bryant, 1972; Bryant, Kovar, and Miller, 1975), a \$10 incentive increased examination cooperation rates relative to no incentive from 70 to 82 percent, but, contrary to expectations, the effect of the promised incentive did not increase as income level decreased. In a parallel remuneration experiment conducted in connection with the second survey (NHANES II) in this series (Findlay and Schaible, 1980), boosting the incentive from \$10 to \$20 increased the overall response rate from 74 to 79 percent, and the increased remuneration was more successful among whites than blacks, but there were no differences by income level. Goetz and his associates (1984) also found no difference in the effectiveness of incentives by education, race (cf. Dohrenwend, 1970), and income.

In contrast, in a nonexperimental comparison, Benus and Ackerman (1971) found that response rates for all major segments of the population were better in a panel where sample members were paid than a comparable panel where sample members were not paid, and they were disproportionately better among low income respondents. Similarly, using nonmonetary incentives in a mail survey, Nederhof (1983) found that incentives produced a disproportionately larger percentage of respondents from lower educational and occupational strata. Though modest, James and Bolstein (1990) found a similar pattern of relationships for level of education and income with various monetary incentives in their

mail survey of cable television subscribers. Ferber and Sudman (1974) also found that monetary incentives were indeed more effective in soliciting cooperation from lower income families than high income households, but such differential effects by education or income level have not been consistently observed in consumer expenditure studies (cf. Ferber, 1974; Cowan, 1977).

In a widely cited but poorly designed study, Gelb (1975) reported a significant difference in the response of lower-class and middle-class respondents to a conditional versus an immediate incentive to return a questionnaire, with middle-class recipients responding better to a prepaid and lower-class recipients to a promised incentive. More recently, Goyder (1990) reported a similar statistical interaction between socioeconomic status (SES) and size and type of incentive, whereby higher SES sample members were more likely to respond to no incentive or a prepaid \$1 incentive, while lower SES sample members were more likely to respond to a post-paid \$10 incentive than high SES sample members.

A more rigorous test of the differential effectiveness of monetary incentives by socioeconomic status was conducted in connection with a field test of 2,000 adults 16 and older for the National Adult Literacy Survey (NALS), a personal interview survey designed to measure one's ability to use printed and written material (Berlin, Mohadjer, Waksberg, Kolstad, Kirsch, Rock, and Yamamoto, 1992). Assessing the impact of incentives of \$0, \$20, and \$35, a significant increase was found in response rates in the two incentive groups (81 and 83 percent, respectively) over the "no incentive" group (73 percent), but no significant difference by incentive level. Of particular importance to the current discussion, they found that incentives were most effective in improving response rates for people with low educational attainment and minority populations. When a monetary incentive was paid, significantly more black and Hispanic adults agreed to take both the background questionnaire and literacy test than when no incentive was offered, and similar results were observed for adults with lower levels of education. Similarly, in a recent pretest for Cycle V of the National Survey of Family Growth (NSFG), a survey of women of childbearing age with an oversample of black and Hispanic women, a \$20 incentive resulted in an overall increase in cooperation of 8 percentage points, but the increase was disproportionately higher for black, Hispanic, and low income women (Duffer, Lessler, Weeks, and Moser, 1994). For example, among lower income women, the \$20 incentive increased response rates by 12 percentage points, three times the level of increase (i.e., 4 percentage points) observed among higher income women. Similarly, the increases observed among black and Hispanic women were 10 and 26 percentage points, respectively, compared to only 2 percentage points for white and other women.

Thus, there is indeed some appreciable evidence that incentives are more likely to influence lower income, lower

socioeconomic subgroups and minority populations to participate in surveys. Overall, however, results regarding the effect of incentives on survey participation from groups differing by race/ethnicity or socioeconomic or status are currently still quite mixed.

3.4 The Conditional Use of Respondent Incentives

The foregoing discussion raises (but clearly doesn't answer) the question of whether all respondents should be paid the same incentive, or should consideration be given to different levels or types of remuneration for different respondents (cf. Peck and Dresch, 1981; Marrett, Kreiger, Dodds, and Hilditch, 1992). A related question raised earlier in our discussion of current practice is "should consideration be given to paying some but not all respondents to a given survey?" In fact, as noted earlier, a number of survey research organizations on some surveys use respondent incentives only for refusal conversion, usually only as a last resort to convert hard-core refusals, the truly hard-to-interview sample members. For example, in the General Social Survey (GSS), conducted by NORC under a grant from the National Science Foundation, "respondent fees" have routinely been used for this purpose; and, in 1989, up front respondent fees were also offered to a select group of respondents who resided in traditionally difficult urban primary sampling units (PSUs), a strategy that resulted in a lowering of the overall cost of the survey and a reduction in the field period, while still achieving the highest completion observed for this survey since 1985 (Law, 1989).

In a paper prepared for the COPAFS symposium in October 1992, Pendleton and Ginsberg (1992) explicitly raised this as an important research question (cf. Tucker, 1992):

[Information is needed on] the most advantageous stage in the data collection process to offer incentives to minimize the cost and time involved in repeated follow up. A comparison of the cost of completed interviews when no incentives are offered, when offering incentives at first contact, and when offering incentives only when response rates are not found to be adequate in terms of cost of completed interview would be useful. (p. 8)

And, although none of them deal with the equity or fairness issues associated with this practice, a few studies have indeed explored the issue of the timing of when respondent incentives are offered, independent of the prepaid versus promised distinction, i.e., the "conditional" use of incentives to convert nonresponders.

For example, in a survey of college undergraduates, Huck and Gleason (1974) found that a follow-up mailing sending a quarter to nonrespondents rather than to everyone on the initial mailout list

cut costs in half, without significantly decreasing the response rate. In essence, the NAEP national assessment experiment with young adults reported by Chromy and Horvitz (1978) was a nonresponse follow-up study. After obtaining a disappointing response rate of only 44 percent with no incentive, the experiment was directed only at nonrespondents from the initial study. While nonrespondents receiving no incentive, but exposed to substantially improved field procedures, were increased to a participation rate of 71 percent, the response rates for those receiving one of three incentive conditions were boosted to 80 to 85 percent. Similarly, in the experiment with young adults surveyed by mail by Peck and Dresch (1981), a prepaid incentive of \$3 in the first wave mailing resulted in a final response rate of 76 percent, compared to only 54 percent among those receiving no incentive; however, a condition in which nonrespondents to the first mailing who initially received no incentive were prepaid \$3 with the second mailing ultimately achieved a response rate of 70 percent. In a household health survey of young adults in Switzerland, Perneger, Etter, and Rougemont (1993) experimented with two "incentives," a promise of 10 Swiss francs (\$7 US) and a red reminder postcard mailed two days after the questionnaire. Initial response rates were 65 percent for those receiving both incentives, 57 percent for those offered the money only, 54 percent for recipients of the reminder card, and 48 percent for those who received neither. Follow-up mailings with incentives sent to all nonrespondents resulted in final response rates of 83, 84, 82, and 78 percent, respectively, attesting to the efficacy of these incentives as a productive nonresponse follow-up device.

In a mail survey of residents of New Zealand with three waves of mailing, a \$1 incentive achieved a final response rate of about 70 percent, regardless of whether it was sent to all respondents with the initial mailing or included in the second mailing for nonrespondents to the first mailout (Brennan, Hoek, and Astridge, 1991). Similarly, in a mail survey of San Diego residents, a monetary incentive of \$5 contingent on response (i.e., a promised incentive) to a second mailing of the questionnaire increased the response rate from initial "nonresponders" by 100 percent relative to controls who received no incentive, and by 75 percent over those who received \$1 not conditional on response (Spry, Hovell, Sallis, Hofstetter, Elder, and Molgaard, 1989).

In a general population telephone survey on family health insurance recently conducted in 10 states for the Robert Wood Johnson Foundation by Mathematica Policy Research, Inc., promised incentives of \$5 and \$10 were compared with no incentive in three states to assess their effects on both response rates and data quality (Strouse and Hall, 1994). Although the \$10 payment marginally increased screener (but not interview) response rates over no incentive prior to refusal conversion, after refusal conversion--which offered a \$10 payment to nonresponders in both groups--cooperation rates for the \$0 and \$10 groups were virtually

equivalent. As a result, payments were retained only for refusal conversion efforts for the remainder of the survey.

Not surprisingly, the potential efficacy of using incentives only for nonresponse follow-up in surveys of physicians has also been examined, with mixed results. Recall that Berry and Kanouse (1987) achieved a 78 percent response rate with a \$20 incentive included with the initial mailing, compared with only 66 percent for those who received a check only after the completed questionnaire was returned. However, a subsample of the original postpayment sample, sent a special follow-up mailing with a check, had a final response rate of 77 percent, indicating that prepayment was effective even if it was used late in the contact process.

In contrast, the recent study reported by Berk, Edwards, and Gay (1993) found that, while "some beneficial impact was found, delaying the incentive until the second round of mailing did not have the same effect as including an incentive with the initial mailing" (p. 241). Overall, a 63 percent response rate was obtained for those receiving a prepaid incentive with the initial mailing, compared with only a 50 percent response rate for those sent a prepaid incentive with a second mailing, and a 40 percent rate for those receiving no incentive.

4. Summary and Conclusion

As we noted in the introduction, the use of monetary incentives and other forms of respondent remuneration has become increasingly common in survey research, spreading rather steadily from the origin of this practice in commercial and market research to the increased use of remuneration in academic and government surveys. In any discussion of this trend among survey research professionals, examples of situations under which incentives are routinely touted as either necessary or highly desirable to achieve adequate response rates always include their use with hard-to-reach or hard-to-interview populations. Although this assertion could be regarded as essentially tautological, there is, in fact, a reasonable consensus on a broad conception of what constitutes "hard-to-reach" respondents, including: (a) hard core refusals; (b) sample members protected by gatekeepers; (c) frequently surveyed groups; (d) target populations or subpopulations among whom the traditional positive forces to cooperate are quite low; and, in general, (e) those who are hard to locate, gain access to, and interview for a variety of different reasons.

In summarizing what we know or think we know about the use of incentives with such populations, it would be most useful to do so in relation to the seven basic questions that we raised at the outset:

1. *Are there indeed specific target populations who are routinely offered remuneration to participate in surveys by most or all survey organizations because they are regarded as especially difficult to survey?*

The specific types of respondents designated as hard-to-reach by organizations that we contacted in preparing this paper include: (1) the economically disadvantaged; (2) the educationally disadvantaged; (3) minority populations; (4) adolescents, youth, and young adults; (5) drug users and those with special health problems; (6) frequently surveyed professional or elite populations; and (7) transients and persons who wish not to be found for legal or other reasons.

Of these seven different categories of respondents, the research literature on the effects of respondent incentives has focused directly on only a few--frequently surveyed professional and elite populations, adolescents and young adults, and the disadvantaged. Overall, to varying degrees, the results of these studies suggest that respondent incentives can be quite effective in stimulating survey cooperation among each of these hard-to-reach populations, especially the first two categories.

2. *While respondent incentives may increase cooperation among initial refusals, are they really effective with hard-core refusals and the truly difficult or impossible to interview populations?*

That respondent incentives are quite effective in averting initial refusals or in converting them after the fact is fairly clear from the research literature, but none of the literature reviewed specifically addresses their efficacy with hard core refusals or those who are truly difficult or impossible to interview. Possible exceptions are the effectiveness of the NALS and NSFG Cycle V field experiments in achieving higher response rates among those subpopulations whose response rates are typically quite low, in spite of extensive follow-up and refusal conversion efforts. In addition, most survey research professionals believe that respondent incentives are an important overall tool in their arsenal for dealing with these difficult-to-interview populations in particular.

3. *Are incentives effective only for certain target populations or subpopulations or more effective for certain population subgroups than for others (i.e., are the effects of incentives different for different population subgroups)?*

While studies of the effectiveness of incentives that focus directly on low income, minority, and disadvantaged populations are quite rare, those that address the more specific question of whether respondent incentives are more effective in improving

cooperation among such target populations than the more affluent and advantaged are considerably more common. Overall, these studies provide some provocative evidence suggesting that incentives are indeed more likely to influence lower income, lower socioeconomic status and minority populations to participate in surveys than those more advantaged. Similarly, several studies of physicians have found significant differences by specialty in the effectiveness of incentives in increasing survey cooperation. More generally, there is a growing body of research evidence suggesting that incentives are more effective for certain populations or population subgroups than for others.

4. *Are incentives really effective in getting past "gatekeepers," either for certain professionals (e.g., physicians) or other difficult-to-survey subgroups of the general population?*

Evidence provided both by survey practitioners and the research literature suggests that respondent incentives can indeed be effective in getting past gatekeepers and gaining access to certain difficult to reach populations.

5. *Are incentives indeed of little use in locating and interviewing hard-to-find cases?*

A few studies and the experiences reported by several research organizations suggest that respondent incentives can indeed be effective in reducing the time and costs associated with locating, tracking, and interviewing highly mobile or otherwise difficult to locate populations.

6. *Should consideration be given to paying some, but not all respondents to a given survey?*

A number of survey research organizations on some surveys use respondent incentives only for refusal conversion, i.e., in general respondents are not paid, but incentives are offered as one tool in the refusal conversion process.

Although there are important questions to be answered related to the equity or fairness of this practice, most of the research literature bearing on its feasibility suggests that quite similar rates of response and response quality can be achieved in most surveys by delaying the use incentives to later stages of contact in the survey process.

7. *Should all respondents be paid the same incentive, or should consideration be given to different levels or types of remuneration for different respondents*

The research literature suggests that incentives (or incentives of a given level) are more effective for certain

population subgroups than for others, and that different respondent groups may be more sensitive to different levels or types (e.g., prepaid versus conditional) of remuneration than others. Thus, some have suggested that "an optimal survey design would probably use different incentives for different population groups" (Peck and Dresch, 1981, p. 256). However, while the observed variation across a number of studies is quite real, the evidence to date does not yet appear sufficient to justify the use of different incentives for different categories of sample members or survey respondents.

Overall, what does this overview suggest about how we define and approach hard-to-reach or hard-to-interview respondents and the relative or special efficacy of remuneration in surveys of these target populations or subpopulations? This question is best addressed perhaps by first approaching the problem from the other direction, i.e., by asking ourselves what motivates cooperative or "easy-to-reach" respondents to participate in surveys.

Fundamentally, it is important to realize that monetary incentives represent only one of many incentives or motivating factors available to survey researchers (cf. Groves, 1992). Survey practitioners already use a number of other types of incentives to encourage survey participation--including appeals to civic duty, the eventual use of the information provided to help people, etc.--and such appeals are generally quite effective. A sense of civic duty to respond to a government survey may be viewed as the cumulative result of the provision of goods and services by the polity to its citizens (Groves, 1992), and this "social contract" to provide information for the public good clearly appeals to many people. Thus, the techniques typically used by survey researchers to encourage participation include appealing to respondents' obligations as citizens and emphasizing the societal benefits likely to accrue from the survey.

In turn, these appeals seek to activate one of two basic types of respondent motive patterns (Cannell and Henson, 1974): (1) a perception that participation in the interview will enable one to achieve certain personal goals (i.e., by emphasizing the importance and purpose of the survey and attempting to link these to achieving some personal or societal goal); and (2) a habitual mode of response toward requests made by legitimate agencies or organizations in society, based on norms of good citizenship, politeness, acquiescence to requests for information, etc. Unfortunately, as Cannell and Henson (1974) note:

these two types of motivations are not very effective in increasing respondent [cooperation or] activity. Respondents do not share the researcher's goals, or, if they do, they fail to see the interview as an effective way of achieving that goal. . . . Similarly, the respondent role may be seen as related to citizen responsibility, but this is not usually a

sufficiently salient or strong motive to induce a high level of activity [or commitment] . . . (p. 313)

In effect, then, an incentive may be offered to respondents to create a personal goal which motivates participation where no other goal or motivation exists.

In this sense, hard-to-reach respondents are those for whom the positive forces to cooperate are quite low, where direct connections to personal goals cannot be readily established and appeals to civic responsibility or benefits to society are quite likely to fail. For example, economically and educationally disadvantaged populations--and others disenfranchised or alienated from the mainstream society--typically have no context for valuing research or their contribution to the research process and are less likely to be persuaded to participate in surveys on the basis of "social utility." It is thus generally unrealistic to expect them to give their time simply because of the potential policy implications of the data we ask them to provide. In short, one must recognize that there are indeed some subpopulations where the "social contract" has broken down, where the dominant mode of decision making is cost-benefit analysis (Groves, 1992) and where economic rather than social exchange processes are paramount.

It is at this intersection that we encounter many of those respondents who we designate as hard-to-reach, and where the use of remuneration takes on major significance as the incentive of choice, as those based on personal goals, obligations to society, and social utility become increasingly blurred or nonexistent. Although these segments of our society may well be growing, they still undoubtedly constitute a clear minority. Thus, a major challenge for those who choose to use respondent incentives to encourage participation among those who are hard-to-survey is to strike a pragmatic and appropriate balance between providing effective levels of remuneration to induce such groups to respond while simultaneously minimizing the risk of alienating those for whom the "social contract" is very much intact, among whom a sense of civic duty is alive and well, thereby continuing to provide a more than adequate basis for their participation in surveys they perceive to be of benefit to themselves, our society, and the nation.

REFERENCES

- Abraham, S. Y. (1994, May). *Using pilot surveys to increase response rates: Findings from the 1992-93 National Study of Postsecondary Faculty*. Paper presented at the 49th Annual Conference of the American Association for Public Opinion Research, Danvers, MA.
- Abraham, S. Y., and Johnson, R. A. (1993, May). *Effects of incentives in a multi-modal survey of elites: Findings from the National Study of Postsecondary Faculty*. Paper presented at the 48th Annual Conference of the American Association for Public Opinion Research, St. Charles, IL.
- Armstrong, J. S. (1975). Monetary incentives in mail surveys. *Public Opinion Quarterly*, 39, 111-116.
- Benus, J., and Ackerman, J. C. (1971). The problem of nonresponse in sample surveys. In J. B. Lansing, S. B. Withey, and A. C. Wolfe (Eds.), *Working papers on survey research in poverty areas* (pp. 26-59). Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan.
- Berk, M. L. (1985). Interviewing physicians: The effect of improved response rate. *American Journal of Public Health*, 75, 1338-1340.
- Berk, M. L., Edwards, W. S., and Gay, N. L. (1993). The use of a prepaid incentive to convert nonresponders on a survey of physicians. *Evaluation and the Health Professions*, 16(2), 239-244.
- Berk, M. L., Mathiowetz, N. A., Ward, E. P., and White, A. A. (1987). The effect of prepaid and promised incentives: Results of a controlled experiment. *Journal of Official Statistics*, 3, 449-457.
- Berlin, M., Mohadjer, L., Waksberg, J., Kolstad, A., Kirsch, I., Rock, D., and Yamamoto, K. (1992). An experiment in monetary incentives. *Proceedings of the Survey Research Methods Section, American Statistical Association*, 393-398.
- Berry, S. H., and Kanouse, D. E. (1987). Physicians response to a mailed survey: An experiment in timing of payment. *Public Opinion Quarterly*, 51, 102-114.
- Blumberg, H. H., Fuller, C., and Hare, A. P. (1974). Response rates in postal surveys. *Public Opinion Quarterly*, 38, 113-123.
- Brennan, M., Hoek, J., and Astridge, C. (1991). The effect of monetary incentives on the response rate and effectiveness of a mail survey. *Journal of the Market Research Society*, 33, 229-241.

Bryant, E. E., Kovar, M. G., and Miller, H. (1975). A study of the effect of remuneration upon response in the Health and Nutrition Examination Survey, United States. *Vital and Health Statistics, Series 2-No.67*. DHEW Pub. No (HRA) 76-1341. Washington, D.C.: U.S. Government Printing Office.

Cannell, C.F., and Camburn, D. (1991, June). *The effect of financial incentives on respondent participation in the Youth Risk Behavior Supplement*. Unpublished manuscript, Survey Research Center, University of Michigan, Ann Arbor, MI.

Cannell, C. F., and Henson, R. (1974). Incentives, motives, and responses bias. *Annals of Economic and Social Measurement*, 3, 307-317.

Cannell, C. F., and Fowler, F. J. (1977). Interviewers and interviewing techniques. In U.S. National Center for Health Services Research. *Advances in health survey research methods: Proceedings of a national invitational conference, 1975*. NCHSR Research Proceedings Series, DHEW Publication No. (HRA) 77-3154. Rockville, MD: Health Resources Administration.

Chromy, J. R., and Horvitz, D. G. (1978). The use of monetary incentives in national assessment household surveys. *Journal of the American Statistical Association*, 73, 473-478.

Church, A. H. (1993). Estimating the effects of incentives on mail response rates: A meta-analysis. *Public Opinion Quarterly*, 57, 62-79.

Cohen, S. B., Walden, D. C., and Ward, E. P. (1992, October). *Providing payments to respondents: The NMES experience*. Paper prepared for the COPAFS/OMB Symposium on Providing Incentives to Survey Respondents, John F. Kennedy School of Government, Harvard University, Cambridge, MA.

Cook, J. R., Schoeps, N., and Kim, S. (1985). Program responses to mail surveys as a function of monetary incentives. *Psychological Reports*, 57, 366.

Council of Professional Associations on Federal Statistics (COPAFS). (1993, September). *Providing incentives to survey respondents: Final report* (Contract No. GS0092AEM0914). Washington, DC: Regulatory Information Service Center, General Services Administration.

Cowan, C. D. (1977). Incentive effects on amounts reported in an expenditure diary survey. *Proceedings of the Social Statistics Section, American Statistical Association*, 498-503.

Dillman, D. A. (1978). *Mail and telephone surveys: The total design method*. New York: John Wiley.

- Dillman, D. A. (1991). The design and administration of mail surveys. *Annual Review of Sociology*, 17, 225-249.
- Dohrenwend, B. S. (1970). An experimental study of payments to respondents. *Public Opinion Quarterly*, 34, 621-624.
- Duncan, W. J. (1979). Mail questionnaires in survey research: A review of response inducement techniques. *Journal of Management*, 5, 39-55.
- Duffer, A. P., Lessler, J. T., Weeks, M. F., and Mosher, W. D. (1994, May). Effects of incentive payments on response rates and field costs in a pretest of a national CAPI survey. Paper presented at the 49th Annual Conference of the American Association for Public Opinion Research, Danvers, MA.
- Erdos, P. L., and Morgan, A. J. (1983). *Professional mail surveys* (rev. ed.). Malabar, FL: Robert E. Krieger Publishing Company.
- Ferber, R. (1974). Consumer expenditure and price data: An overview. *Annals of Economic and Social Measurement*, 3, 299-305.
- Ferber, R., and Sudman, S. (1974). Effects of compensation in consumer expenditure studies. *Annals of Economic and Social Measurement*, 3, 319-331.
- Findlay, J. S., and Schaible, W. L. (1980). A study of the effect of increased remuneration on response in a health and nutrition examination survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 590-594.
- Fox, R. J., Crask, M. R., and Kim, J. (1988). Mail survey response rate: A meta-analysis of selected techniques for inducing response. *Public Opinion Quarterly*, 52, 467-491.
- Furse, D. H., and Stewart, D. W. (1982). Monetary incentives versus promised contribution to charity: New evidence on mail survey response. *Journal of Marketing Research*, 19, 375-380.
- Furse, D. H., and Stewart, D. W. (1984). Manipulating dissonance to improve mail survey response. *Psychology and Marketing*, 1 (2), 79-94.
- Gelb, B. D. (1975). Incentives to increase survey returns: Social class considerations. *Journal of Marketing Research*, 12, 107-109.
- Godwin, R. K. (1979). The consequences of large monetary incentives in mail surveys of elites. *Public Opinion Quarterly*, 43, 378-387.

Goetz, E. G., Tyler, T. R., and Cook, F. L. (1984). Promised incentives in media research: A look at data quality, sample representativeness, and response rate. *Journal of Marketing Research*, 21, 148-154.

Goodman, L. and Jensen, L. (1981). Economic surveys of medical practice: AMA's periodic survey of physicians, 1966-1978. In S. Sudman (Ed.), *Proceedings of the third biennial conference on health survey methods*. Rockville, MD: National Center for Health Services Research.

Goodstadt, M. S., Chung, L., Kronitz, R., and Cook, J. (1977). Mail survey response rates: Their manipulation and impact. *Journal of Marketing Research*, 14, 391-395.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, 25, 161-178.

Goyder, J. C. (1990). *Cash incentives and socio-economic bias in surveys*. Unpublished manuscript, Department of Sociology, University of Waterloo, Ontario, Canada.

Groves, R. M. (1989). *Survey errors and survey costs*. New York: John Wiley and Sons.

Groves, R. M. (1992, October). *Respondent incentives, minimum mean square error per unit cost, and the resident's responsibility to the polity*. Discussion notes prepared for the COPAFS/OMB Symposium on Providing Incentives to Survey Respondents, John F. Kennedy School of Government, Harvard University, Cambridge, MA.

Groves, R. M., Cialdini, R. B., and Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56, 475-495.

Gunn, W. J., and Rhodes, I. N. (1981). Physician response rates to a telephone survey: Effects of monetary incentive level. *Public Opinion Quarterly*, 45, 109-115.

Hackler, J. C., and Bourgette, P. (1973). Dollars, dissonance and survey returns. *Public Opinion Quarterly*, 37, 276-281.

Hansen, R. A. (1980). A self-perception interpretation of the effect of monetary and nonmonetary incentives on mail survey respondent behavior. *Journal of Marketing Research*, 17, 77-83.

Harvey, L. (1987). Factors affecting response rates to mailed questionnaires: A comprehensive literature review. *Journal of the Market Research Society*, 29, 341-353.

- Heberlein, T. A., and Baumgartner, R. (1978). Factors affecting response rates to mailed questionnaires: A quantitative review of the published literature. *American Sociological Review*, 43, 447-462.
- Hopkins, K. D., and Gullickson, A. R. (1992). Response rates in survey research: A meta-analysis of the effects of monetary gratuities. *Journal of Experimental Education*, 61(1), 52-62.
- Hopkins, K. D., Hopkins, B. R., and Schon, I. (1988). Mail surveys of professional populations: The effects of monetary gratuities on return rates. *Journal of Experimental Education*, 56(3), 173-175.
- Houston, M. J., and Ford, N. M. (1976). Broadening the scope of methodological research on mail surveys. *Journal of Marketing Research*, 13, 397-403.
- Hubbard, R., and Little, E. L. (1988a). Promised contributions to charity and mail survey responses: Replication with extension. *Public Opinion Quarterly*, 52, 223-230.
- Hubbard, R., and Little, E. L. (1988b). Cash prizes and mail survey response rates: A threshold analysis. *Journal of the Academy of Marketing Science*, 16, 42-44.
- Huck, S. W., and Gleason, E. M. (1974). Using monetary inducements to increase response rates from mail surveys. *Journal of Applied Psychology*, 59, 222-225.
- James, J. M., and Bolstein, R. (1990). The effect of monetary incentives and follow-up mailings on the response rate and response quality in mail surveys. *Public Opinion Quarterly*, 54, 346-361.
- James, J. M., and Bolstein, R. (1992). Large monetary incentives and their effect on mail response rates. *Public Opinion Quarterly*, 56, 442-453.
- Jobber, D., and Saunders, J. (1988). Modelling the effects of prepaid monetary incentives on mail-survey response. *Journal of the Operational Research Society*, 39(4), 365-372.
- Kanuk, L., and Berenson, C. (1975). Mail surveys and response rates: A literature review. *Journal of Marketing Research*, 12, 440-453.
- Kephart, W. M., and Bressler, M. (1958). Increasing the response to mail questionnaires. *Public Opinion Quarterly*, 22, 123-132.

Kerachsky, S. H., and Mallar, C. D. (1981). The effects of monetary payments on survey responses: Evidence from a longitudinal study of economically disadvantaged youths. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 258-263.

Kulka, R. A. (1992, October). A brief review of the use of monetary incentives in federal statistical surveys. Paper prepared for the COPAFS/OMB Symposium on Providing Incentives to Survey Respondents, John F. Kennedy School of Government, Harvard University, Cambridge, MA.

Law, J. W. (1989). *The impact of respondent fee payments on the 1989 General Social Survey*. Chicago: National Opinion Research Center, University of Chicago.

Linsky, A. S. (1975). Stimulating responses to mailed questionnaires: A review. *Public Opinion Quarterly*, 39, 82-101.

Lockhart, D. C. (1991). Mailed surveys to physicians: The effect of incentives and length on the return rate. *Journal of Pharmaceutical Marketing and Management*, 6(1), 107-121.

Marrett, L. D., Kreiger, N., Dobbs, L., and Hilditch, S. (1992). The effect on response rates of offering a small incentive with a mailed questionnaire. *Annals of Epidemiology*, 2(5), 745-753.

Martin, B. (1984). *Don't survey physicians*. Chicago: American Medical Association.

McDaniel, S. W., and Rao, C. P. (1980). The effect of monetary inducement on mailed questionnaire response quality. *Journal of Marketing Research*, 17, 265-268.

Miller, H. W., Kennedy, J., and Bryant, E. E. (1972). A study of the effect of remuneration upon response in a health and nutrition examination survey. *Proceedings of the Social Statistics Section, American Statistical Association*, 1972, 370-375.

Mizes, J. S., Fleece, E. L., and Roos, C. (1984). Incentives for increasing return rates: Magnitude levels, response bias, and format. *Public Opinion Quarterly*, 48, 794-800.

Mooney, G., Giesbrecht, L., and Shettle, C. (1993). To pay or not to pay; That is the question. Paper presented at the 48th Annual Conference of the American Association for Public Opinion Research, St. Charles, Illinois, May 20-23.

Nederhof, A. J. (1983). The effects of material incentives in mail surveys: Two studies. *Public Opinion Quarterly*, 47, 103-111.

- Paolillo, J. G. P., and Lorenzi, P. (1984). Monetary incentives and mail questionnaire response rates. *Journal of Advertising*, 13(1), 46-48.
- Peck, J. K., and Dresch, S. P. (1981). Financial incentives, survey response, and sample representativeness: Does money matter? *Review of Public Data Use*, 9, 245-266.
- Pendleton, A., and Ginsburg, A. (1992, October). *Providing incentives to survey respondents in education program evaluations*. Paper prepared for the COPAFS/OMB Symposium on Providing Incentives to Survey Respondents, John F. Kennedy School of Government, Harvard University, Cambridge, MA.
- Perneger, T. V., Etter, J-F., and Rougemont, A. (1993). Randomized trial use of a monetary incentive and a reminder card to increase the response rate to a mailed health survey. *American Journal of Epidemiology*, 138, 714-722.
- Robin, D. P., and Walters, C. G. (1976). The effect on return rate of messages explaining monetary incentives in mail questionnaire studies. *Journal of Business Communications*, 13, 49-54.
- Shuttleworth, F. K. (1931). A study of questionnaire technique. *Journal of Educational Psychology*, 22, 652-658.
- Sudman, S. (1985). Mail surveys of reluctant professionals. *Evaluation Review*, 9, 349-360.
- Sudman, S., and Ferber, R. (1974). A comparison of alternative procedures for collecting consumer expenditure data for frequently purchased products. *Journal of Marketing Research*, 11, 128-135.
- Skinner, S. J., Ferrell, O. C., and Pride, W. M. (1984). Personal versus nonpersonal incentives in mail surveys: Immediate versus delayed inducements. *Journal of the Academy of Marketing Science*, 12, 106-114.
- Spry, V. M., Hovell, M. F., Sallis, J. G., Hofstetter, C. R., Elder, J. P., and Molgaard, C. A. (1989). Recruiting survey respondents to mailed surveys: Controlled trials of incentives and prompts. *American Journal of Epidemiology*, 130(1), 166-172.
- Strouse, R., and Hall, J. (1994, May). *Effects of payments on response rates and data quality for a general population telephone survey*. Paper presented at the 49th Annual Conference of the American Association for Public Opinion Research, Danvers, MA.

Tambor, E. S., Chase, G. A., Faden, R. R., Geller, G., Hofman, K. J., and Holtzman, N. A. (1993). Improving response rates through incentive and follow-up: The effects on a survey of physicians' knowledge of genetics. *American Journal of Public Health*, 83, 1599-1603.

Tucker, C. (1992, October). *Providing incentives to respondents in government surveys*. Paper prepared for the COPAFS/OMB Symposium on Providing Incentives to Survey Respondents, John F. Kennedy School of Government, Harvard University, Cambridge, MA.

Weber, S. J., Wycoff, M. L., and Adamson, D. R. (1982). *The impact of two clinical trials on physician knowledge and practice*. Arlington, VA: Market Facts, Inc.

Weinberg, E. (1973). The issue of payment to respondents--wise or unwise. *Public Opinion Quarterly*, 37 (3), 478-480.

Weiss, C. H. (1975). Interviewing in evaluation research. In E.I. Struening and M. Guttentag (Eds.), *Handbook of evaluation research* (pp. 355-395). Beverly Hills, CA: Sage Publishing Company.

Willimack, D. K. (1993). Selected results of the incentive experiment on the 1992 Farm Costs and Returns Survey. *Proceedings of the International Conference on Establishment Surveys* (pp. 490-495). Alexandria, VA: American Statistical Association.

Willimack, D. K., Petrella, M., Beebe, T., and Welk, M. (1992, August). *The use of incentives in surveys: Annotated bibliography*. Survey Research Center, Institute for Social Research, University of Michigan.

Wotruba, T. R. (1966). Monetary inducements and mail questionnaire response. *Journal of Marketing Research*, 3, 398-400.

Yammarino, F. J., Skinner, S. J., and Childers, T. L. (1991). Understanding mail survey response behavior: A meta-analysis. *Public Opinion Quarterly*, 55, 613-639.

Yu, J., and Cooper, H. (1983). A quantitative review of research design effects on response rates to questionnaires. *Journal of Marketing Research*, 20, 36-44.

Zusman, B. J., and Duby, P. (1987). An evaluation of the use of monetary incentives in postsecondary survey research. *Journal of Research and Development in Education*, 20(4), 73-78.

Summary of Resolvable Errors in Edit Output

Benzene

chkcode codable primary occupation (qprmy_cd) - housewife (aqid=03233);
freqs qisfx2 = 0 for 1 case
rtickicd look into the operative procedure codes on page 1 - should these codes
 be based on residual body site?
 page 61, aqid=01187: can this be coded per earlier error resolution

TCE

mainedit flag155-multibirth no, 1 preg, 2 livebirths- aqid=06011
chkcode 1 case w/emp_cd=10- do not assign- aqid=08222
phone area codes- page 1 -obs 29 and 30: 910 should 919
 duplicate phone #:
 area code=815; qid=009020090-check last name-it appears it may be
 the same person who was a proxy with same first and middle names
 but different last name
 addresses with more than 1 phone:
 251 ...- fix 6th digit in 1st phone #
 4329.. - put last digit in 3rd position for first observation

street page 69: blocks 5 and 6
 page 88 blocks 4 and 5
 page 126 - last 2 blocks

Transaction log

- 1) The transaction log should be log of CHANGES MADE to database - not changes INTENDED to be made - describe how this log is created with respect to the SAS program where corrections are programmed.
- 2) provide report of errors - changes not made- as provided earlier.
- 3) local variables (i.e. AQID=2795: LREGZIP) should never be changed
- 4) need consistent handling of #'s(i.e. AQID=0118: "#10"), apt, NO (1st vs. 2nd line of address)
- 5) re-sort this report: aqid, variable name, reverse chronological date and time
- 6) appears to be 2 qids for aqid=2923. Please look into this and also provide proc freq/list of aqid * qid.
- 7) NORC needs to do very careful QA/QC of transaction log prior to finalizing files and sending final deliverable - perhaps forward this to ATSDR for review electronically (scrambled) prior to review of other deliverables

specific problems:

<u>AQID</u>	<u>Problem</u>
0118	"Court" should be standardized
0192	"Calendar, CT" appears in final edit output, but in this log seems to be an old value that was changed to "Venice, CA". How can this be?
0802	Cancer condition (2/91) appears to have occurred BEFORE last interview date
2692	Variable Name blank
2876	PK should be PARK
2923	*2 qids?? why so many changes - is this correct? need details on case
3954	condition (D/89) appears to have occurred BEFORE last interview date

DISCUSSION

Diane K. Willimack
National Agricultural Statistics Service

I would like to begin by congratulating and commending these authors on two thorough and well-presented papers. Kulka has written a theory section that sets the stage for incentive use in surveys and raises conceptual issues for consideration by survey designers. He goes on to provide a thorough review of the literature and a thoughtful evaluation of the implications for survey designers, particularly regarding "hard-to-reach" population subgroups.

Ezzati, White, Mosher, and Sanchez -- the NCHS authors plus one -- have provided a compilation of the findings from a series of experiments in which incentive use has been systematically studied on a wide variety of survey types and target populations -- from hospitals to households, from physicians to disadvantaged population subgroups. In addition they report the value to interviewers of incentive use: how respondent incentives enhance interviewers' confidence and their ability to do their job, as well as how interviewer incentives enhance their morale and motivation. This summary challenges us all to consider how to apply their findings to our own survey situations, so that we need not reinvent the wheel.

Allow me to share what I learned from these two papers. Kulka differentiates the social norm of reciprocity and social exchange theory from the theory of economic exchange as the conceptual basis for incentive use in surveys. Typically we describe incentives by their type -- monetary vs. nonmonetary -- and timing -- prepaid vs. promised. Let us focus on timing. Use of prepaid incentives is based on the social theories, while promised incentives are believed to invoke economic exchange in respondents.

These authors tend to use the terms "incentive" and "remuneration" interchangeably. Kulka also uses the word "compensation". I decided to look up these words in my Random House Collegiate Dictionary (1988). I know that our use of words as jargon need not have any relationship with the English language, but I thought this exercise might be instructive, as well as help me clarify my thoughts.

"Incentive," in my dictionary, is defined as "something that incites to action," and lists as synonyms, "stimulus, spur, incitement, encouragement." Let me interpret: encouragement to respond. "Remuneration" is "something that remunerates; reward; pay." "Remunerate" means "to recompense," which means "to repay or reward (someone), as for aid or service." Again I offer my interpretation: to recompense (repay) for aiding us by responding to our survey. Finally, "compensation" is "something given or received as an equivalent for services, etc.," with synonyms "recompense, remuneration, payment." It seems to me that "remuneration" and "compensation" are interchangeable with each other, but not with "incentive."

Let us return to theory. One of the problems with promised incentives is that, in many studies comparing them to prepaid incentives, the amount of money or the item was token in nature. Thus while the "promise" tends to invoke economic exchange in respondents, the token nature was not sufficient to influence respondents to engage in the economic contract. But remuneration does

invoke economic exchange, since it implies an amount of money sufficiently close to the value of the service of survey response rendered by the respondent.

Thus, if an incentive is token in nature, it must be prepaid, for it can only rely upon social norms in order to be successful. If a promised incentive is to be used, then, for it to be successful, it must be substantial enough to be considered remuneration or compensation in an economic exchange. Of course the latter must take into consideration the difficulty of the burden of the task of completing the survey. Perhaps I have just exhibited a firm grasp of the obvious, but this notion did not crystalize for me until I read these two papers side by side.

However, tokens appear to work in instances when signs of appreciation and good faith are meaningful to the respondent. For example, the National Agricultural Statistics Service (NASS) conducted an incentive experiment on the 1992 Farm Costs and Returns Survey, an annual voluntary survey of U.S. farm operators collecting detailed expenditure and income data in personal interviews lasting 90 minutes on average. Response rates were increased by nearly 5.5 percentage points in the group receiving a prepaid nonmonetary incentive. Of greater interest, though, was the finding that incentive use increased response rates by 17 points among the smallest farms (those with sales less than \$20,000) and by nearly 12 points among the largest farms with sales of \$500,000 or more. The token pocket portfolio and calculator that we gave them cannot have indicated an economic exchange to these large farms. Instead, it likely had appeal as a symbol of appreciation consistent with the repeated survey contact to which these two groups, in particular, are subjected (Willimack, 1993).

Farms are establishments, and NASS surveys collect establishment information from them. Although farms may be a special case because they exhibit many of the same characteristics as households, application of incentives or remuneration in establishment surveys is not trivial. The NCHS paper reports a survey of hospital records, in which hospitals are paid for their service of abstracting sampled records. This is clearly remuneration with basis in economic exchange. However this may not always be so clear cut for incentive use or remuneration in establishment surveys, particularly surveys of businesses.

Incentives are meant to "incite action," to influence the decision to participate in a survey and to motivate the respondent. But in an establishment survey, it is unclear who we are attempting to influence with incentive use. In an establishment, the person making the decision about survey participation may not be the desired respondent. The desired respondent is the person who is the most knowledgeable provider of the information being sought, usually the person who has access and understanding of any records to be used as a source for responding (Edwards and Cantor, 1991).

If a prepaid incentive is provided in an establishment survey, who gets it? Is it the boss, in order to "incite" or encourage a favorable decision about responding? Or is it the employee who actually completes the survey instrument, in order to motivate careful response? Or is it the business, as income or as payment for the service of completing the survey? It does seem less difficult to make this "leap" if it is remuneration being offered rather than an incentive. Remuneration is more clearly a payment to the business for the service of completing the questionnaire. But then the dollar amount must be reasonable relative to the burden of the task.

Let us consider further the dollar amount to be used as an incentive or as an offer of remuneration. I think money is tricky, unless the amount is clearly token. Money is a sensitive topic for many people, and it has all kinds of different connotations. No matter how much is offered, there will always be someone for whom that amount is not enough. Furthermore, since respondents use all the information available to them when responding to survey questions, it seems reasonable to suggest that the same is true in those initial moments of contact during which a survey participation decision is being made. A key piece of information is the dollar amount being offered. I think the dollar amount provides an indication to respondents about the potential difficulty of the survey task, and may arouse suspicion. The amounts of \$20, \$30, \$50, even \$100 in several of the health studies described by the NCHS authors may have provided a fair indication of the difficulty of the task. On the other hand, consider the James and Bolstein (1992) study cited by Kulka in which a promise of \$50 failed to increase response rates in a survey of construction subcontractors. Similarly, in the HIV survey described in the NCHS paper, \$175 offered to survey refusers increased response rates by only 4 percentage points, while \$100 resulted in a 10 point increase. These inordinately large dollar amounts may have indicated a difficult survey task to the respective target populations, resulting in a sense, "Well, if you're offering that much money, then what you're asking me to do must be really hard, embarrassing, or uncomfortable. It must be something that I don't want to do." Meanwhile, the choice to accept or reject the offer -- to engage in the economic contract -- remains with the respondent.

The amount of money offered as an incentive or as remuneration is an important variable to the survey designer. Here we have reached the edge of a gaping hole in the literature on incentives. What is the trade-off between the incentive or remuneration amount and the difficulty of the survey task? Indeed, under which survey circumstances will a token incentive suffice, and when is remuneration needed? These papers seem to suggest that when the survey task is particularly intrusive, burdensome, lengthy, or longitudinal, then the promise of remuneration seems appropriate, if not necessary. But these factors must be evaluated relative to the expectations and perceptions of the members of the target population: Perhaps a personal interview lasting 1½ to 3 hours and requesting detailed expense and income data is considered by farm operators to be as personal and sensitive as the Survey of Family Growth asking women aged 15-44 about abortion and sexual practices. As discussants are required to say, "more research is needed."

Furthermore, participants at the 1992 COPAFS conference on incentives listed guidelines for OMB to consider when evaluating incentive use in Federal surveys (COPAFS, 1993). These include:

"To compensate a respondent if there is risk in participating."

"When there are unusual demands or intrusions on the respondent (e.g., lengthy interviews, keeping a diary, having a blood sample drawn, ...)."

"When sensitive questions are being asked."

"If there is a lengthy field period (e.g., a commitment over time for a panel survey)."

"If there is any out-of-pocket cost to the respondent ..."

"If the respondent is a small business or a nonprofit institution in a voluntary survey and the respondent perceives some cost and burden to participating."

Although these represent only a subset of COPAF's recommendations to OMB, it appears to me that they outline survey circumstances that favor remuneration, as I have defined it in this discussion. However, there may be survey situations in which an incentive, as I have differentiated it, will suffice. If OMB policy appears to favor remuneration over incentive use, then agencies will design surveys accordingly, with consequences for survey budgets, management, and respondent burden. OMB needs to write policy that recognizes and encourages appropriate use of both incentives and remuneration.

How do we evaluate the effectiveness of incentives or remuneration in surveys? In the early days, as Kulka points out in his literature review, it was as simple as "Were response rates increased by incentive use?" A resounding "YES" is supported by a large number of studies on various types of surveys. Soon attention turned to the quality of the data, adding the question, "Does incentive use improve the quality of the data through the respondent's increased attentiveness to the survey task, particularly reflected in reduced item nonresponse?", or, at least, "Can we be sure that data quality has not been reduced?" Again, both Kulka and the NCHS authors provide a great deal of evidence supporting improved data quality related to incentive use or remuneration.

Let me next turn attention to what I call potential "dividends" to be achieved from incentive use: reduction in components of total survey error. Both sets of authors provide a number of examples of increased survey participation among "hard-to-reach" or typically under-represented population subgroups. In addition, the NCHS authors have provided results that show systematic differences on key variables among groups whose participation was increased by remuneration. Thus, these papers present evidence that nonresponse bias in survey estimates may be reduced through use of incentives or remuneration.

Moreover, let me offer an additional "dividend" found in two incentive experiments in which I have been involved: enhanced ability to identify ineligible sample units, resulting in nonsampling error reduction. In an incentive experiment on the 1991 Detroit Area Study, a statistically significant increase was found in the rate of ineligible sample addresses. This may have been due, in part, to more diligent postal returns of the small packages containing a prepaid pen incentive, along with more reliable interviewer confirmation, when an address was a vacant housing unit, a business, or an incorrect address (Willimack and others, 1994). A replication of this result is currently being tested in the 1994 Detroit Area Study.

In addition, in the incentive experiment conducted on the 1992 Farm Costs and Returns survey, a higher rate of sample units being screened out as non-farms, and therefore not eligible for the survey, was found to be statistically significant. Incentive recipients who had no agriculture may have been more attentive to the survey request and more determined to notify the interviewer of their non-farm status, rather than to refuse or be inaccessible based on a belief that the survey did not apply to them (Willimack, 1993).

Incentive studies of the 1990's appear to have added cost evaluations or cost-benefit analyses of incentive use or remuneration. The results of the pretest of the Survey of Family Growth are particularly compelling. They show that the \$20 incentive/remuneration actually saved money over

no incentive, since these respondents required fewer contacts and were less likely to break appointments, resulting in reduced interviewing costs.

Cost evaluation is an area that still needs development. Achieved cost savings may not always be so clear. Instead, frequently we are left to judge whether observed benefits were sufficient to justify the additional costs of incentive/remuneration use. It is difficult to put a dollar value on increased response rates, improved data quality, and reduction in total survey error. No doubt we all consider these to be very valuable. But how much improvement is needed in order to justify the additional cost of incentives or remuneration? What production and efficiency measures should be monitored during data collection and post-survey review and processing? What comparisons with which other survey methods should be undertaken? We must be able to evaluate whether incentive use or remuneration is the best tool to pull out of our methodology tool box.

Again I congratulate the authors, and I thank them for their contributions to our understanding of incentive use in surveys.

REFERENCES

Council of Professional Associations on Federal Statistics (COPAFS) (1993), Providing Incentives to Survey Respondents: Final Report (Contract No. GS0092AEM0914), Washington, D.C.: Regulatory Information Service Center, General Services Administration.

Edwards, W. S., and Cantor, D. (1991), "Toward a Response Model in Establishment Surveys," Ch. 12 in Measurement Errors in Surveys, edited by P. P. Beimer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, and S. Sudman, New York: John Wiley and Sons, Inc., pp. 211-233.

James, J. M., and Bolstein, R. (1992), "Large Monetary Incentives and Their Effect on Mail Survey Response Rates," Public Opinion Quarterly 56:442-453.

The Random House Collegiate Dictionary (1988), New York: Random House, Inc.

Willimack, D. K. (1993), "Selected Results of the Incentive Experiment on the 1992 Farm Costs and Returns Survey," Proceedings of the Interantional Conference on Establishment Surveys, Alexandria, VA: Amaerican Statistical Association, pp. 490-495.

Willimack, D. K., Schuman, H., Pennell, B. E., and Lepkowski, J. M. (1994), "The Effects of a Prepaid Nonmonetary Incentive on Response Rates and Response Quality in a Face-to-Face Survey," unpublished manuscript, Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI.

DISCUSSION

W. Sherman Edwards
Westat, Inc.

The use of monetary incentives in surveys is an endlessly fascinating topic for those of us in the business of collecting and analyzing survey data. Regardless of the importance and policy relevance of any particular survey, regardless of the eminent history and usefulness of the statistics derived from a survey, at some point in data collection it gets down to the trenches -- getting the last reluctant respondent to cooperate, finding the last elusive subject, persuading the last indifferent citizen to answer what seem to him like senseless and repetitive questions. The use of monetary incentives to help in this unavoidable struggle raises interesting methodological questions and, even more provocatively, questions about our relationship with our respondents. As a society, we are strongly ambivalent about money; it is the most public of our life's trappings, and yet at the same time it is the most private. Every survey researcher knows that it is not questions about sexual behavior or the intimate details of health that elicit the highest item refusal rates -- it is questions about income and financial assets. Introducing money into any relationship that is not primarily an economic one complicates that relationship, as anyone who has loaned money to a relative or friend can attest.

My discussion is organized first around two specific methodological questions and then addresses the relationship issue. Although the literature on the effects of monetary incentives is extensive, I will suggest several areas that warrant considerable further study.

The first methodological question is the most basic, "Do monetary incentives improve response rates?" The two papers today provide a fairly unequivocal "Yes" response to this question for surveys where incentive experiments were conducted. The paper by Dick Kulka examines the use of incentives for hard-to-survey populations. This is a very comprehensive paper; it provides a thorough and very useful overview of the research literature on incentives, of different theoretical views of how incentives work, and of the methodological issues raised when considering monetary incentives. The paper then goes on to explore in more detail issues around incentives for hard-to-survey populations. Although this comment is not particularly relevant to my discussion, I can't help but note that I never before considered the commonality among drug users, physicians, youth in general, CEOs, impoverished urban minorities, diabetics, unwed mothers, small farm operators, and those defaulting on student loans. (They are all considered hard to survey, and thus worthy of consideration for monetary incentives.)

The paper by Trena Ezzati and her colleagues reviews the experiences with monetary incentives of several surveys about health and the use of health care services conducted by the National Center for Health Statistics and the Agency for Health Care Policy and Research. This paper presents a more detailed look at a smaller number of studies, including some that have pushed at the edges of the envelope containing this issue.

Kulka surveyed participants in the 1993 COPAFS conference on incentives, and found their use widespread. I think it's fair to characterize the view of many survey firms towards

incentives as one of a number of tools that they can use to achieve satisfactorily high response rates, but one that is often used with some reluctance. Other tools in this kit, of course, include repeated callbacks to convert initially reluctant respondents, the use of specially designed persuasive materials that may be targeted to particular reasons for refusal, and the use of especially effective interviewers. How do we choose among these tools in designing our data collection strategies? The most sympathetic view is that we choose rationally, based on empirical evidence and careful consideration of all alternatives, choosing the most cost-effective procedures possible. A more cynical view might be that many surveys or survey organizations opt for incentives because they are easier to implement than other approaches. I suspect that both views have some truth. For mail surveys, a fair amount of thought and research have been devoted to optimal design strategies, most notably Dillman's Total Design Method (Dillman, 1978). The literature on telephone surveys is somewhat less well developed, and that on in-person still less. This is in part because the issues become more complex as the relationship between survey and respondent becomes more personal. Recent work by Groves and Couper (1994, e.g.) is notable in applying constructs and research techniques from several disciplines to the tool kit. One lesson from their work is that we still have a lot to learn about the tools, including when and how to use them.

An important part of the question of whether incentives work is whether they are cost-effective. Some of the research studies cited in the two papers included formal evaluations of cost-effectiveness. The recent NSFG evaluation, for example, found that the cost per case for the \$20 incentive treatment was lower than for the non-incentive treatment, in addition to yielding higher response rates. Other studies have examined the cost of monetary incentives against other methods, such as repeated mailings, to achieve comparable response rates. Often, incentives seem to reduce the cost of surveys. The cost-effectiveness of a monetary incentive is related to the mode of administration and the size of the incentive. The marginal cost of a \$20 incentive for an in-person survey is clearly much lower than the cost of a similar incentive in a mail survey. Kulka mentions the possible effectiveness of an incentive in locating elusive respondents. In-person locating is one of the most expensive of data collection activities -- even a small improvement in locating effectiveness would probably be worth the cost of a monetary incentive judiciously mentioned to an informant.

The second methodological question is, "What effect do monetary incentives have on data quality?" Kulka notes two commonly expressed views from the research literature: the first that respondents view incentives as a kind of "social exchange," and thus work harder and provide better responses. The second is that incentives are an "external motivator," decreasing internal motivation and hence reducing the quality of responses. Here, the literature is somewhat less persuasive. Some studies, like the NSFG experiment and the seroprevalence survey pilot cited by Ezzati, take the "more is better" view, which is probably often appropriate, particularly when the results are as dramatic as in these studies. Few studies have combined incentive experiments with validation of data, which is often difficult or even impossible. In particular, how does data quality differ between those who would (or do) participate without an incentive and those for whom the incentive makes the difference? To examine this, we might use respondents' perceptions about the use of incentives as an explanatory variable for some measures of data quality.

The only study cited that explicitly explored respondents' attitudes was associated with the NHIS Youth Risk Behavior Supplement, as cited by Kulka. Kalton and his colleagues used follow-up cognitive interviews with youths responding to a field test of the survey to discover an apparently strong link between the incentive and the effort expended in answering questions. This seems an especially promising methodology for exploring the effects of incentives, particularly on "special" or hard-to-survey populations.

Monetary incentives are frequently used in diary surveys -- such use is even sanctioned by the Office of Management and Budget. Could a monetary incentive affect the behavior of interest by a diary keeper? At the recent AAPOR conference, Diane Woodard of Arbitron (1994) described a series of incentive experiments for respondents in their radio listenership survey. One treatment involved a sweepstakes with a \$3,000 first prize in addition to the usual nominal monetary incentive. In this treatment, all radio ratings increased. The post hoc explanation of this phenomenon was that respondents had increased their radio listening (or at least their reporting) in the erroneous belief that this would increase their chances of winning the sweepstakes. Another explanation, of course, is that the sweepstakes respondents were simply reporting better, although I am inclined to accept Woodard's interpretation.

A strongly held view of many survey researchers is that all respondents should be subject to exactly the same stimulus, as nearly as this can be controlled by the survey design. Only in this way can we be confident of the reliability of survey responses. This view has come under increasing criticism from those who view the structured interview as too restrictive a vehicle for meaningful exchange of information. For the more traditional view, monetary incentives may throw a monkey wrench in the works. As Kulka notes, some surveys offer incentives only to "hard-core" refusals or to those who are expected to be difficult to locate or interview. Even when the same incentive is offered to all respondents, their reactions to being "paid" for their time undoubtedly vary considerably. In every survey I've been associated with that offered incentives, some respondents refused to accept them. Again, it would seem useful to ask respondents how they feel about the payment, and to examine how responses vary by these views. At the heart of the matter, do monetary incentives introduce more variation in respondent motivation and perception of the survey than otherwise exists?

Let me now turn to the issue of the relationship between the survey and the respondent. If we define this relationship in terms of the use of incentives, there seem to be three views of the nature of the relationship. The first, which I would associate with not using incentives, is that surveys are a social good, and that participation alone is sufficient reward. Clearly, government surveys are the most logical candidates for this high-minded view, and there is some evidence in support of it. The Census Bureau does not use incentives, and yet achieves the highest response rates in the industry, even setting aside their mandatory surveys.

Those of us in the private sector, however, tend to be fairly pragmatic, if still somewhat ambivalent, as Kulka reports, about the use of incentives. What claim does Westat or anyone else outside of the government have on respondents, even if we do often represent the government? (I note that all of the surveys described in the Ezzati et al paper were conducted by private contractors.) Even if we have ethical concerns about the fairness of paying respondents in some surveys but not others, or of paying the "hard-core" but not the compliant, are the

alternatives any more attractive? Is it ethically more appealing to ask again and again, or to fly in the charming interviewer from Indiana, than to offer a monetary incentive?

Another view of the relationship between survey and respondent is that of the "social exchange" that Kulka describes. I would suggest that many who conduct survey research are most comfortable with this view. I have heard many researchers, some in describing incentive experiments, say, "The interviewers feel better when they can give the respondent something." Interviewers in this case, as they so often are, are our proxies. Each of today's papers had one mention of the positive effect of respondent incentives on interviewer behavior. This is another area that warrants further study.

One of the groundbreaking studies described by Ezzati et al is the pilot test of the National Household HIV Seroprevalence Survey, conducted by RTI in Dallas. Here, the size of the incentives offered (\$50, \$100, \$175) seem to go well beyond what would be called for in a social exchange view of the interview. Respondents were asked in return for a blood sample and to complete a brief sexual history questionnaire. However, the context of this survey was socially charged. Respondents might be fearful of the blood draw in general, but particularly because it was associated with an AIDS survey. One view of the size of the incentive was that it was an appropriately attractive exchange for asking respondents to overcome a whole set of fears.

The third view of monetary incentives in this relationship is that of a business transaction, or "economic exchange," as it is described in the Kulka paper -- we are paying the respondent for his or her time. The implication of this view is that we are participants in an information marketplace, subject to the principles of supply and demand. In government household surveys, even those conducted by contractors, most researchers shy away from this view, and even explicitly deny it. Even some incentive experiments wind up giving all respondents the same payment. The survey of allergists cited by Ezzati is an example -- regardless of the incentive treatment, all participating physicians were eventually paid the same amount.

One is really taking an economic view when one uses disproportionate payments for different survey respondents, including payment to initial nonresponders but not to compliers. Here, we pay according to how badly we want the information -- truly supply and demand. This practice is anathema to some researchers, who cite the ethics of fairness and their distaste for rewarding noncompliant behavior.

Where monetary incentives have sometimes become a business transaction is, not surprisingly, in surveys of businesses. The prediction about respondents coming to expect incentives has probably come to pass for physician surveys, although Kulka notes that some are still conducted successfully without incentives. Ezzati et al describe the experience of the National Hospital Discharge Survey, where hospitals are routinely paid piecework for completing survey forms. The National Medical Expenditure Survey's Medical Provider Survey also allowed payment for survey respondents, the amount to be negotiated, but only if requested by the medical provider. On the other hand, most Federal establishment surveys do not offer reimbursement on demand or incentives.

One argument put forward against monetary incentives is that respondents will come to expect payment, and response rates will be even harder to achieve at a reasonable cost. This is referred to in the summary of last year's COPAFS symposium as the "slippery slope" argument. In essence, this argument says that the use of incentives breeds the perception of a survey marketplace. It seems unlikely that this effect would occur among the general public solely from Federal surveys. Even with the current rate of surveying, most households are never or rarely selected for a government survey (excepting the decennial Census, of course). If we are in or to be in a survey marketplace in this sense, commercial surveys will have had more to do with that result, and OMB has no control over their behavior. The "slippery slope" is more likely to refer to the behavior of survey organizations if we come to rely too routinely on monetary incentives to achieve our response rate goals.

One noticeable feature of both of these well-prepared and very interesting papers is the number of experiments that have been conducted around incentives, particularly for Federal surveys. I suspect that as we in the statistical community continue to explore the complexities of our relationship with our survey respondents, we will continue to conduct incentive experiments at an unflagging pace. There is certainly much more to learn about the effects of incentives on data quality, and about alternative methods for influencing respondent compliance.

REFERENCES

Dillman, D.A. (1978), Mail and Telephone Surveys: The Total Design Method, New York: Wiley.

Groves, R. and Couper, M. (1994), "Understanding the Process of Survey Cooperation: Household-Interviewer Doorstep Interactions," paper presented at the 49th Annual Conference of the American Association of Public Opinion Research, Danvers, MA.

Woodard, D. (1994), "Raising Response Rates with Premiums, or Is \$1 Worth Anything to Anyone Anymore?" paper presented at the 49th Annual Conference of the American Association of Public Opinion Research, Danvers, MA.